PCT/US 2004/010582

REC'D 22 JUN 2005

WIPO PCT

P2 1320744

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

June 17, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY OF THE BELOW IDENTIFIED INTERNATIONAL APPLICATION AS ORIGINALLY FILED AND ANY CORRECTIONS THERETO FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE ACTING AS A RECEIVING OFFICE UNDER THE PATENT COOPERATION TREATY.

APPLICATION NUMBER: ***PCT/US04/06308***
FILING DATE: ***March 02, 2004***

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

H. L. Jackson

H. L. JACKSON
Certifying Officer

DT02 Rec'd PCT/PTO 0 2 MAR 2004

PCT/US04/06308

TRANSMITTAL LETTER TO THE UNITED STATES RECEIVING OFFICE

Date	2 March 2004
International Application No.	Not Yet Assigned PCT/US 04/06308
Attorney Docket No.	PROL-PWO-024

I. Certificate under 37 CFR 1.10 (if applicable)

EV323 524 199US	2 March 2004
Express Mail mailing number	Date of Deposit

I hereby certify that the application/correspondence attached hereto is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to MS PCT, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Maura A. Gallagher	Maura A. Gallagher
Signature of person mailing correspondence	Typed or printed name of person mailing correspondence

II. [X] **New International Application**

TITLE	POSH INTERACTING PROTEINS AND RELATED METHODS	Earliest priority date (Day/Month/Year) 03/03/03

SCREENING DISCLOSURE INFORMATION: In order to assist in screening the accompanying international application for purposes of determining whether a license for foreign transmittal should and could be granted and for other purposes, the following information is supplied. (Note: check as many boxes as apply):

A. [ ] The invention disclosed was not made in the United States.

B. [ ] There is no prior U.S. application relating to this invention.

C. [X] The following prior U.S. application(s) contain subject matter which is related to the invention disclosed in the attached international application. *(NOTE: priority to these applications may or may not be claimed on form PCT/RO/101 (Request) and this listing does not constitute a claim for priority.)*

App No		App No	60/475,825 filed 3 June 2003
App No	60/451,437 filed 3 March 2003	App No	60/479,317 filed 17 June 2003
App No	60/452,284 filed 5 March 2003	App No	60/480,376 filed 19 June 2003
App No	60/456,640 filed 20 March 2003	App No	60/480,215 filed 19 June 2003
App No	60/460,526 filed 3 April 2003	App No	60/493,860 filed 8 August 2003
App No	60/464,285 filed 21 April 2003	App No	60/503,931 filed 16 September 2003
App No	60/469,462 filed 9 May 2003	App No	60/455,760 filed 19 March 2003
App No	60/471,378 filed 15 May 2003	App No	60/460,792 filed 4 April 2003
App No	60/472, 327 filed 20 May 2003	App No	60/498,634 filed 28 August 2003
App No	60/474,706 filed 30 May 2003	App No	US03/35712 filed 10 November 2003
App No	A PCT application filed on February 5, 2004 (Attorney Docket No. PROL-PWO-039), in the name of Iris Alroy, Daniel Taglicht, Yuval Reiss, Liora Yaar, and Shmuel Tuvia entitled "Posh Associated Kinases and Related Methods."	App No	A provisional application filed on March 2, 2004, (Attorney Docket No. PROL-P79-024), in the name of Daniel N. Taglicht, Iris Alroy, Yuval Reiss, Liora Yaar, Danny Ben-Avraham, Shmuel Tuvia and Tsvika Greener entitled "Posh Interacting Proteins and Related Methods."

D. [X] The present international application [X] contains additional subject matter not found in the prior U.S. application(s) identified in paragraph C. above. The additional subject matter is found on pages THROUGHOUT and [ ] **DOES NOT ALTER** [X] **MIGHT BE CONSIDERED TO ALTER** the general nature of the invention in a manner which would require the U.S. application to have been made available for inspection by the appropriate defense agencies under 35 U.S.C. 181 and 37 CFR 5.1. See 37 CFR 5.15.

IV. [ ] **A Request for Rectification under PCT Rule 91** [ ] **A Petition** [ ] **A Sequence Listing Diskette**

V. [X] Other (please specify): Request & Fee Calculation Sheet (7 pp); Description (155 pp); Claims (16 pp); Abstract (1 p); Drawings (202 pp); Return postcard from RO/US confirming receipt of PCT application

The person signing this form is the:		
	[ ] Applicant	Kathleen Ehrhard Typed name of signer
	[X] Attorney/Agent Reg. No. P-55,144	Kathleen Ehrhard Signature
	[ ] Common Representative	

PCT/US04/06308

HOME COPY

PCT

REQUEST

The undersigned requests that the present international application be processed according to the Patent Cooperation Treaty.

For receiving Office use only

International Application No. PCT/US 04/06308

International Filing Date 02 MAR 2004 (02.03.04)

PCT INTERNATIONAL APPLICATION RO/US

Name of receiving Office and "PCT International Application"

Applicant's or agent's file reference (if desired) (12 characters maximum) PROL-PWO-024

Box No. I TITLE OF INVENTION

POSH INTERACTING PROTEINS AND RELATED METHODS

Box No. II APPLICANT ☐ This person is also inventor

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

PROTEOLOGICS, INC.
40 Ramland Road South
Suite 10
Orangeburg, New York 10962
United States of America

Telephone No.

Facsimile No.

Teleprinter No.

Applicant's registration No. with the Office

State *(that is, country)* of nationality: US

State *(that is, country)* of residence: US

This person is applicant for the purposes of: ☐ all designated States ☒ all designated States except the United States of America ☐ the United States of America only ☐ the States indicated in the Supplemental Box

Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

TAGLICHT, Daniel N.
Lapid
Israel

This person is:

☐ applicant only

☒ applicant and inventor

☐ inventor only *(If this check-box is marked, do not fill in below.)*

Applicant's registration No. with the Office

State *(that is, country)* of nationality: IL

State *(that is, country)* of residence: IL

This person is applicant for the purposes of: ☐ all designated States ☐ all designated States except the United States of America ☒ the United States of America only ☐ the States indicated in the Supplemental Box

☒ Further applicants and/or (further) inventors are indicated on a continuation sheet.

Box No. IV AGENT OR COMMON REPRESENTATIVE; OR ADDRESS FOR CORRESPONDENCE

The person identified below is hereby/has been appointed to act on behalf of the applicant(s) before the competent International Authorities as: ☒ agent ☐ common representative

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)*

EHRHARD, Kathleen
Ropes & Gray LLP
One International Place
Boston, Massachusetts 02110-2624
United States of America

Telephone No. (617) 951-7037

Facsimile No. (617) 951-7050

Teleprinter No.

Agent's registration No. with the Office P-55,144

☐ **Address for correspondence:** Mark this check-box where no agent or common representative is/has been appointed and the space above is used instead to indicate a special address to which correspondence should be sent.

PCT/US04/06308

Continuation of Box No. III **FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)**

If none of the following sub-boxes is used, this sheet should not be included in the request.

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

ALROY, Iris
Hashirion Street 10/17
74065 Nes Ziona
Israel

This person is:
- [ ] applicant only
- [x] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

Applicant's registration No. with the Office

State *(that is, country)* of nationality: IL

State *(that is, country)* of residence: IL

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [x] the United States of America only
- [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

REISS, Yuval
Hahavazelet 11/6
Kiriat-ono
Israel

This person is:
- [ ] applicant only
- [x] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

Applicant's registration No. with the Office

State *(that is, country)* of nationality: IL

State *(that is, country)* of residence: IL

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [x] the United States of America only
- [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

YAAR, Liora
8 Kalisher Street
43354 Raanana
Israel

This person is:
- [ ] applicant only
- [x] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

Applicant's registration No. with the Office

State *(that is, country)* of nationality: IL

State *(that is, country)* of residence: IL

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [x] the United States of America only
- [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

BEN-AVRAHAM, Danny
Igal Alon 20
Zichron Jackov
Israel

This person is:
- [ ] applicant only
- [x] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

Applicant's registration No. with the Office

State *(that is, country)* of nationality: IL

State *(that is, country)* of residence: IL

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [x] the United States of America only
- [ ] the States indicated in the Supplemental Box

- [x] Further applicants and/or (further) inventors are indicated on another continuation sheet.

PCT/US04/06308

Continuation of Box No. III **FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)**

If none of the following sub-boxes is used, this sheet should not be included in the request.

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (i.e. country) of residence if no State of residence is indicated below.)*

TUVIA, Shmuel
Hartzit 1
42490 Netanya
Israel

This person is:
- [ ] applicant only
- [x] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

Applicant's registration No. with the Office

State *(that is, country)* of nationality: IL

State *(that is, country)* of residence: IL

This person is applicant for the purposes of: [ ] all designated States [ ] all designated States except the United States of America [x] the United States of America only [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

GREENER, Tsvika
Hahavazelet 9a
Ness Ziona
Israel

This person is:
- [ ] applicant only
- [x] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

Applicant's registration No. with the Office

State *(that is, country)* of nationality: IL

State *(that is, country)* of residence: IL

This person is applicant for the purposes of: [ ] all designated States [ ] all designated States except the United States of America [x] the United States of America only [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

This person is:
- [ ] applicant only
- [ ] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

Applicant's registration No. with the Office

State *(that is, country)* of nationality:

State *(that is, country)* of residence:

This person is applicant for the purposes of: [ ] all designated States [ ] all designated States except the United States of America [ ] the United States of America only [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

This person is:
- [ ] applicant only
- [ ] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

Applicant's registration No. with the Office

State *(that is, country)* of nationality:

State *(that is, country)* of residence:

This person is applicant for the purposes of: [ ] all designated States [ ] all designated States except the United States of America [ ] the United States of America only [ ] the States indicated in the Supplemental Box

[ ] Further applicants and/or (further) inventors are indicated on another continuation sheet.

Supplemental Box *If the Supplemental Box is not used, this sheet should not be included in the request.*

1. *If, in any of the Boxes except Boxes Nos. VIII(i) to (v) for which a special continuation box is provided, the space is insufficient to furnish all the information: in such case, write "Continuation of box No. ...." (indicate the number of the Box) and furnish the information in the same manner as required according to the captions of the Box in which the space was insufficient, in particular:*

   Application No *if more than two persons are to be indicated as applicants and/or inventors and no "continuation sheet" isavailable: in such case, write "Continuation of Box No. III" and indicate for each additional person the same type of information as required in Box No. III. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below;*

   (ii) *if, in Box No. II or in any of the sub-boxes of Box No. III, the indication "the States indicated in the Supplemental Box" is checked: in such case, write "Continuation of Box No. II" or "Continuation of Box No. III" or "Continuation of Boxes No. II and No. III" (as the case may be), indicate the name of the applicant(s) involved and, next to (each) such name, the State(s) (and/or, where applicable, ARIPO, Eurasian, European or OAPI patent) for the purposes of which the named person is applicant;*

   (iii) *if, in Box No. II or in any of the sub-boxes of Box No. III, the inventor or the inventor/applicant is not inventor for the purposes of all designated States or for the purposes of the United States of America: in such case, write "Continuation of Box No. II" or "Continuation of Box No. III" or "Continuation of Boxes No. II and No. III" (as the case may be), indicate the name of the inventor(s) and, next to (each) such name, the State(s) (and/or, where applicable, ARIPO, Eurasian, European or OAPI patent) for the purposes of which the named person is inventor;*

   (iv) *if, in addition to the agent(s) indicated in Box No. IV, there are further agents: in such case, write "Continuation of Box No. IV" and indicate for each further agent the same type of information as required in Box IV;*

   (v) *if, in Box No. VI, there are more than three earlier applications whose priority is claimed: in such case, write "Continuation of Box No. VI" and indicate for each additional earlier application the same type of information as required in Box No. VI.*

2. *If, the applicant intends to make an indication of the wish that the international application be treated, in certain designated States, as an application for a patent of addition, certificate of addition, inventor's certificate of addition or utility certificate of addition: in such a case, write the name or two-letter code of each designated States concerned and the indication "patent of addition," "certificate of addition," "inventor's certificate of addition" or " utility certificate of addition," the number of the parent application or parent patent or other parent grant and the date of grant of the parent patent or other patent grant or the date of filing of the parent application (Rules 4.11(a)(iii) and 49bis.1(a) or (b)).*

3. *If the applicant intends to make an indication of the wish that the international application be treated, in the United States of America, as a continuation or continuation-in-part of an earlier application: in such a case, write "United States of America" or "US" and the indication "continuation" or "continuation-in-part" and the number and the filing date of the parent application (Rules 4.11(a)(iv) and 49bis.1(d)).*

Continuation of Box No. VI

(4) Date: 05 March 2003 (05/03/03) Application: 60/452284
National Application Country: US
(5) Date: 20 March 2003 (20/03/03) Application: 60/456640
National Application Country: US
(6) Date: 03 April 2003 (03/04/03) Application: 60/460526
National Application Country: US
(7) Date: 21 April 2003 (21/04/03) Application: 60/464285
National Application Country: US
(8) Date: 15 May 2003 (15/05/03) Application: 60/471378
National Application Country: US
(9) Date: 20 May 2003 (20/05/03) Application: 60/472327
National Application Country: US
(10) Date: 30 May 2003 (30/05/03) Application: 60/474706
National Application Country: US
(11) Date: 03 June 2003 (03/06/03) Application: 60/475825
National Application Country: US
(12) Date: 17 June 2003 (17/06/03) Application: 60/479317
National Application Country: US
(13) Date: 19 June 2003 (19/06/03) Application: 60/480215
National Application Country: US
(14) Date: 08 August 2003 (08/08/03) Application: 60/493860
National Application Country: US
(15) Date: 16 September 2003 (16/09/03) Application: 60/503931
National Application Country: US
(16) Date: 02 March 2004 (02/03/04) Application: A provisional application filed on March 2, 2004, (Attorney Docket No. PROL-P79-024), in the name of Daniel N. Taglicht, Iris Alroy, Yuval Reiss, Liora Yaar, Danny Ben-Avraham, Shmuel Tuvia and Tsvika Greener entitled "Posh Interacting Proteins and Related Methods."
National Application Country: US
(17) Date: 03 March 2003 (03/03/03) Application: 60/451,437
National Application Country: US
(18) Date: 09 May 2003 (09/05/03) Application: 60/469,462
National Application Country: US
(19) Date: 19 June 2003 (19/06/03) Application: 60/480,376
National Application Country: US
(20) Date: 10 November 2003 (10/11/03) Application: US03135712
National Application Country: US
(21) Date: PCT filed 05 February 2004 (Attorney Docket No. PROL-PWO-039) in the name of Iris Alroy, Daniel Taglicht, Yuval Reiss, Liora Yaar, and Shmuel Tuvia entitled "Posh Associated Kinases and Related Methods"

Continuation of Box No. IV:

Steven Baglio, 51,426; J. Steven Baughman, 47,414; Mark W. Bellomy, 51,452; John V. Bianco, 36,748; Johnny Y. Chen, 46,614; James P. Demers, 34,320; Gojeb L. Frehywot, 52,916; Gloria Fuentes, 47,580; Gregory Glover, 34,173; William G. Gosz, 27,787; Patricia Granahan, 32,227; Z. Angela Guo, 54,144; David P. Halstead, 44,735; Margaret E. Jamroz, 54,196; Edward J. Kelly, 38,936; Charles Larsen, 48,533; Agnes S. Lee, 46,862; Paul E. Lewkowicz, 44,870; Weishi Li, 53,217; Yu Lu, 50,306; Alexander Manganiello, 53,264; Robert A. Mazzarese, 42,852; Christopher Natkanski, 50,365; R. Daniel O'Connor, P54,343; Ignacio Perez de la Cruz, 55,535; Melissa S. Rones, Ph.D., 54,408; Spencer H. Schneider, 45,923; Sanjay Sitlani, 48,489; Wolfgang E. Stutius, 40,256; Erika Takeuchi, 55,661; Lisa Treannie, 41,368; Anita Varma, 43,221; Matthew P. Vincent, 36,709; Dalila Argaez Wendlandt, 52,351; and Levina Wong, P54,551

And all other agents of:
ROPES & GRAY LLP, Patent Group
One International Place
Boston, Massachusetts 02110-2624
United States of America
Customer ID No: 28,120

Box No. V DESIGNATIONS

The filing of this request constitutes under Rule 4.9(a), the designation of all Contracting States bound by the PCT on the international filing date, for the grant of every kind of protection available and, where applicable, for the grant of both regional and national patents.

However,

☐ DE Germany is not designated for any kind of national protection.

☐ KR Republic of Korea is not designated for any kind of national protection.

☐ RU Russian Federation is not designated for any kind of national protection.

(The check-boxes above may be used to exclude (irrevocably) the designations concerned in order to avoid the ceasing of the effect, under the national law, of an earlier national application from which priority is claimed. See the Notes to Box No. V as to the consequences of such national law provisions in these and certain other States.)

Box No. VI PRIORITY CLAIM

The priority of the following earlier application(s) is hereby claimed:

	Filing date of earlier application *(day/month/year)*	Number of earlier application	Where earlier application is: national application: country or Member of WTO	regional application:* regional Office	international application receiving Office
item (1)	19 March 2003 (19.03.2003)	60/455760	US		
item (2)	28 August 2003 (28.08.2003)	60/498634	US		
item (3)	04 April 2003 (04.04.2003)	60/460792	US		

☒ Further priority claims are indicated in the Supplemental Box.

The receiving Office is requested to prepare and transmit to the International Bureau a certified copy of the earlier application(s) *(only if the earlier application was filed with the Office which for the purposes of this international application is the receiving Office)* identified above as:

☒ all items ☐ item (1) ☐ item (2) ☐ item (3) ☒ other, see Supplemental Box

* *Where the earlier application is an ARIPO application, indicate at least one country party to the Paris Convention for the Protection of Industrial Property or one Member of the World Trade Organization for which that earlier application was filed (Rule 4.10(b)(ii)):*

Box No. VII INTERNATIONAL SEARCHING AUTHORITY

Choice of International Searching Authority (ISA) *(if two or more International Searching Authorities are competent to carry out the international search, indicate the Authority chosen; the two-letter code maybe used):*

ISA /US

Request to use results of earlier search; reference to that search *(if an earlier search has been carried out by or requested from the International Searching Authority):*

Date *(day/month/year)* Number Country *(or regional Office)*

Box No. VIII DECLARATIONS

The following declarations are contained in Boxes Nos. VIII (i) to (v) *(mark the applicable check-boxes below and indicate in the right column the number of each type of declaration):*

Number of declarations

☐ Box No. VIII (i) Declaration as to the identity of the inventor :

☐ Box No. VIII (ii) Declaration as to the applicant's entitlement, as at the international filing date, to apply for and be granted a patent :

☐ Box No. VIII (iii) Declaration as to the applicant's entitlement, as at the international filing date, to claim the priority of the earlier application :

☐ Box No. VIII (iv) Declaration of inventorship (only for the purposes of the designation of the United States of America :

☐ Box No. VIII (v) Declaration as to non-prejudicial disclosures or exceptions to lack of novelty :

PCT/US04/06308

Box No. IX CHECK LIST; LANGUAGE OF FILING

This international application contains:

(a) **in paper form**, the following number of sheets:

request (including declaration sheets)	:	6
description (excluding sequence listings and/or tables related thereto)	:	155
claims	:	16
abstract	:	1
drawings	:	202
Sub-total number of sheets	:	
sequence listings	:	
tables related thereto	:	
(for both, actual number of sheets if filed in paper form, whether or not also filed in computer readable form; see (c) below)		
Total number of sheets	:	380

(b) ☐ **only in computer readable form** (Section 801(a)(i))
- (i) ☐ sequence listings
- (ii) ☐ tables related thereto

(c) ☐ **also in computer readable form** (Section 801(a)(ii))
- (i) ☐ sequence listings
- (ii) ☐ tables related thereto

Type and number of carriers (diskette, CD-ROM, CD-R or other) on which are contained the
- ☐ sequence listing:
- ☐ tables related thereto:

(additional copies to be indicated under items 9(ii) and/or 10(ii), in right column)

This international application is **accompanied by the following** item(s) *(mark the applicable check-boxes below and indicate in right column the number of each item):*

	Item	Number of items
1.	☒ fee calculation sheet	: 1
2.	☐ original separate power of attorney	:
3.	☐ original general power of attorney	:
4.	☐ copy of general power of attorney; reference number, if any: --	:
5.	☐ statement explaining lack of signature	:
6.	☐ priority document(s) identified in Box No. VI as item(s):	:
7.	☐ translation of international application into *(language)*:	:
8.	☐ separate indications concerning deposited microorganisms or other biological material	:
9.	☐ sequence listing in computer readable form *(indicate type and number of carriers)*	
	(i) ☐ copy submitted for the purposes of international search under Rule 13*ter* only (and not as part of the international application)	:
	(ii) ☐ *(only where check-box (b)(i) or (c)(i) is marked in left column)* additional copies including, where applicable, the copy for the purposes of international search under Rule 13*ter*	:
	(iii) ☐ together with relevant statement as to the identity of the copy or copies with the sequence listings part mentioned in left column	:
10.	tables in computer readable form related to sequence listing *(indicate type and number of carriers)*	
	(i) ☐ copy submitted for the purposes of international search under Section 802 (*b-quater*) only (and not as part of the international application)	:
	(ii) ☐ *(only where check-box (b)(ii) or (c)(ii) is marked in left column)* additional copies including, where applicable, the copy for the purposes of international search under Section 802 (*b-quater*)	:
	(iii) ☐ together with relevant statement as to the identity of the copy or copies with the tables mentioned in left column	:
11.	X other *(specify)*: Return postcard from RO/US confirming receipt of PCT & encls.	:

Figure of the drawings which should accompany the abstract:

Language of filing of the international application: English

Box No. X SIGNATURE OF APPLICANT, AGENT OR COMMON REPRESENTATIVE

Next to each signature, indicate the name of the person signing and the capacity in which the person signs (if such capacity is not obvious from reading the request).

Kathleen Ehrhard

Kathleen Ehrhard, Agent for Applicant(s)
ROPES & GRAY LLP

02.03.04

For receiving Office use only

1.	Date of actual receipt of the purported international application:	DT02 Rec'd PCT/PTO 02 MAR 2004
3.	Corrected date of actual receipt due to later but timely received papers or drawings completing the purported international application:	
4.	Date of timely receipt of the required corrections under PCT Article 11(2):	
5.	International Searching Authority (if two or more are competent:	ISA / US
6.	☐ Transmittal of search copy delayed until search fee is paid.	

2. Drawings:
- ☑ received:
- ☐ not received:

For International Bureau use only

Date of receipt of the record copy by the International Bureau:

PCT/US04/06308

This sheet is not part of and does not count as a sheet of the international application.

PCT

FEE CALCULATION SHEET
Annex to the Request

For receiving Office use only

International Application No. PCT/US 04/06308 (02 03 04)

Date stamp of the receiving Office 02 MAR 2004

Applicant's or agent's file reference: PROL-PWO-024

Applicant: Proteologics, Inc., et al.

CALCULATION OF PRESCRIBED FEES

			For receiving Office use only
1. TRANSMITTAL FEE	300.00	T	300-
2. SEARCH FEE	1,000.00	S	1000-

International search to be carried out by US

(If two or more International Searching Authorities are competent to carry out the international search, indicate the name of the Authority which is chosen to carry out the international search)

3. INTERNATIONAL FILING FEE

Where item (b) and/or (c) of Box No. IX apply, enter Sub-total number of sheets
Where item (b) and (c) of Box No. IX do not apply, enter Total number of sheets } 380

i1	first 30 sheets	1,035.00	i1	1035
i2	350 (number of sheets in excess of 30) x 11.00 (fee per sheet) =	3,850.00	i2	3850
i3	additional component (only if sequence listing and /or tables related thereto are filed in computer readable form under Section 801(a)(i), or both in that form and on paper, under Section 801(a)(ii)): 400 x ______ (fee per sheet) =		i3	

Add amounts entered at i1, i2 and i3 and enter total at I	4,885.00	I	4885

(Applicants from certain States are entitled to a reduction of 75% of the international fee. Where the applicant is (or all applicants are) so entitled, the total to be entered at I is 25% of the international filing fee.)

4. FEE FOR PRIORITY DOCUMENT *(if applicable)*	320.00	P	420-
5. TOTAL FEES PAYABLE	$ 6,505.00	TOTAL	6605-

Add amounts entered at T, S, I and P, and enter total in the TOTAL box

MODE OF PAYMENT

- [x] authorization to charge deposit account (see below)
- [ ] postal money order
- [ ] cash
- [ ] coupons
- [ ] cheque
- [ ] bank draft
- [ ] revenue stamps
- [ ] other *(specify)*:

AUTHORIZATION TO CHARGE (OR CREDIT) DEPOSIT ACCOUNT

(This mode of payment may not be available at all receiving Offices)

- [x] Authorization to charge the total fees indicated above.
- [x] *(This check-box may be marked only if the conditions for deposit accounts of the receiving office so permit)* Authorization to charge any deficiency or credit any overpayment in the total fees indicated above.
- [x] Authorization to charge the fee for priority document.

Receiving Office: RO/ US

Deposit Account No.: 18-1945

Date: 02 March 2004

Name: Kathleen Ehrhard

Signature: Kathleen Ehrhard

Form PCT/RO/101 (Annex) (March 2001)

See Notes to the fee calculation sheet

9382278.1

POSH INTERACTING PROTEINS AND RELATED METHODS

RELATED APPLICATIONS

This application claims the benefit of priority of U.S. Provisional Application number 60/451,437 filed 3 March 2003; 60/452,284 filed 5 March 2003; 60/456,640 filed 20 March 2003; 60/460,526 filed 3 April 2003; 60/464,285 filed 21 April 2003; 60/469,462 filed 9 May 2003; 60/471,378 filed 15 May 2003; 60/472,327 filed 20 May 2003; 60/474,706 filed 30 May 2003; 60/475,825 filed 3 June 2003; 60/479,317 filed 17 June 2003; 60/480,376 filed 19 June 2003; 60/480,215 filed 19 June 2003; 60/493,860 filed 8 August 2003; 60/503,931 filed 16 September 2003; 60/455,760 filed 19 March 2003; 60/460,792 filed 4 April 2003; 60/498,634 filed 28 August 2003; and a provisional application filed on March 2, 2004, (Attorney Docket No. PROL-P79-024), in the name of Daniel N. Taglicht, Iris Alroy, Yuval Reiss, Liora Yaar, Danny Ben-Avraham, Shmuel Tuvia, and Tsvika Greener entitled "Posh Interacting Proteins and Related Methods"; a PCT application US03/35712 filed 10 November 2003; and a PCT application filed on February 5, 2004, (Attorney Docket No. PROL-PWO-039), in the name of Iris Alroy, Daniel Taglicht, Yuval Reiss, Liora Yaar, and Shmuel Tuvia entitled "Posh Associated Kinases and Related Methods". The teachings of the referenced Applications are incorporated herein by reference in their entirety.

BACKGROUND

Potential drug target validation involves determining whether a DNA, RNA or protein molecule is implicated in a disease process and is therefore a suitable target for development of new therapeutic drugs. Drug discovery, the process by which bioactive compounds are identified and characterized, is a critical step in the development of new treatments for human diseases. The landscape of drug discovery has changed dramatically due to the genomics revolution. DNA and protein sequences are yielding a host of new drug targets and an enormous amount of associated information.

The identification of genes and proteins involved in various disease states or key biological processes, such as inflammation and immune response, is a vital part

of the drug design process. Many diseases and disorders could be treated or prevented by decreasing the expression of one or more genes involved in the molecular etiology of the condition if the appropriate molecular target could be identified and appropriate antagonists developed. For example, cancer, in which one or more cellular oncogenes become activated and result in the unchecked progression of cell cycle processes, could be treated by antagonizing appropriate cell cycle control genes. Furthermore many human genetic diseases, such as Huntington's disease, and certain prion conditions, which are influenced by both genetic and epigenetic factors, result from the inappropriate activity of a polypeptide as opposed to the complete loss of its function. Accordingly, antagonizing the aberrant function of such mutant genes would provide a means of treatment. Additionally, infectious diseases such as HIV have been successfully treated with molecular antagonists targeted to specific essential retroviral proteins such as HIV protease or reverse transcriptase. Drug therapy strategies for treating such diseases and disorders have frequently employed molecular antagonists which target the polypeptide product of the disease gene(s). However, the discovery of relevant gene or protein targets is often difficult and time consuming.

One area of particular interest is the identification of host genes and proteins that are co-opted by viruses during the viral life cycle. The serious and incurable nature of many viral diseases, coupled with the high rate of mutations found in many viruses, makes the identification of antiviral agents a high priority for the improvement of world health. Genes and proteins involved in a viral life cycle are also appealing as a subject for investigation because such genes and proteins will typically have additional activities in the host cell and may play a role in other non-viral disease states.

Other areas of interest include the identification of genes and proteins involved in cancer, apoptosis and neural disorders (particularly those associated with apoptotic neurons, such as Alzheimer's disease).

It would be beneficial to identify proteins involved in one or more of these processs for use in, among other things, drug screening methods. Additionally, once a protein involved in one or more processes of interest has been identified, it is possible to identify proteins that associate, directly or indirectly, with the initially

identified protein. Knowledge of interactors will provide insight into protein assemblages and pathways that participate in disease processes, and in many cases an interacting protein will have desirable properties for the targeting of therapeutics. In some cases, an interacting protein will already be known as a drug target, but in a different biological context. Thus, by identifying a suite of proteins that interact with an initially identified protein, it is possible to identify novel drug targets and new uses for previously known therapeutics.

SUMMARY

This application provides isolated, purified or recombinant complexes comprising a POSH polypeptide and one or more POSH-associated protein (POSH-AP). In certain aspects, the POSH-AP comprises a polypeptide selected from the group consisting of: PKA, SNX1, SNX3, ATP6V0C, PTPN12, PPP1CA, GOSR2, CENTB1, DDEF1, ARF1, ARF5, PACS-1, EPS8L2, HERPUD1, UNC84B, MSTP028, GOCAP, EIF3S3, SRA1, CBL-B, RALA, SIAH1, SMN1, SMN2, SYNE1, TTC3, VCY2IP1 and UBE2N (UBC13). In other aspects, the POSH-AP comprises a polypeptide selected from the group consisting of: ARHV (Chp), WASF1, HIP55, SPG20, HLA-A, and HLA-B. In further aspects, the POSH-AP comprises one or more polypeptides set forth in Table 8. In certain embodiments the POSH polypeptide is a human POSH polypeptide.

In certain embodiments, this application provides isolated, purified or recombinant complexes comprising a HERPUD1 polypeptides and a ubiquitin ligase, examples of the ubiquitin ligase include CBL-B, TTC3, and SIAH1.

In certain embodiments, the application provides methods for identifying agents that modulates an activity of a POSH polypeptide or POSH-AP, comprising identifying an agent that disrupts a complex of a POSH polypeptide and a POSH-AP, wherein an agent that disrupts such a complex is an agent that modulates an activity of the POSH polypeptide or the POSH-AP.

In yet other embodiments, the application provides methods of identifying an antiviral agent, comprising identifying a test agent that disrupts a complex comprising a POSH polypeptide and a POSH-AP and evaluating the effect of the test agent on either a pro-infective or pro-replicative function of a virus is an

antiviral agent, wherein an agent inhibits such a function of a virus is an antiviral agent. In certain embodiments the POSH-AP is selected from the group consisting of: PKA, SNX1, SNX3, PTPN12, GOSR2, CENTB1, ARF1, ARF5, PACS-1, EPS8L2, HERPUD1, SMN1, SMN2, UNC84B, MSTP028, GOCAP, CBL-B, SYNE1, UBE2N (UBC13), SIAH1, TTC3, WASF1, HIP55, RALA, and SPG20. Examples of such viruses include for example, envelope viruses such as the Human Immunodeficiency Virus, the West Nile Virus, and the Moloney Murine Leukemia Virus (MMuLV).

In other embodiments, the application provides methods of identifying an anti-apoptotic agent, comprising identifying a test agent that disrupts a complex comprising a POSH polypeptide and a POSH-AP and evaluating the effect of the test agent on apoptosis of a cell wherein an agent that decreases apoptosis of the cell is an anti-apoptotic agent. In yet other embodiments, the application provides methods of identifying an anti-cancer agent, comprising identifying a test agent that disrupts a complex comprising a POSH polypeptide and a POSH-AP and evaluating the effect of the test agent on proliferation or survival of a cancer cell, wherein an agent that decreases proliferation or survival of a cancer cell is an anti-cancer agent. Examples of the POSH-AP include PKA, SNX1, PTPN12, PPP1CA, ARF1, ARF5, CENTB1, EPS8L2, EIF3S3, CBL-B, RALA, SIAH1, TTC3, ATP6V0C, and VCY2IP1. In certain embodiments, the cancer is a POSH-associated cancer.

In certain aspects, the application provides methods of identifying an agent that inhibits trafficking of a protein through the secretory pathway, comprising identifying a test agent that disrupts a complex comprising a POSH polypeptide and a POSH-AP and evaluating the effect of the test agent on the trafficking of a protein through the secretory pathway wherein an agent that disrupts localization of said POSH-AP is an agent that inhibits trafficking of a protein through the secretory pathway. In certain embodiments, the protein is a myristoylated protein. In yet other embodiments, the protein is a viral protein. In alternative embodiments, the protein is associated with a neurological disorder such as for example the amyloid beta precursor protein.

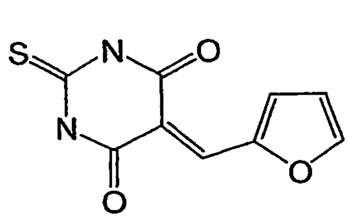
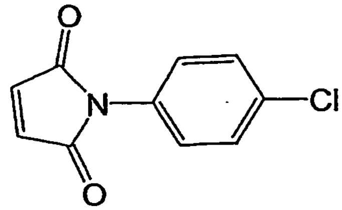
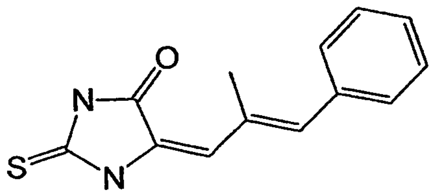
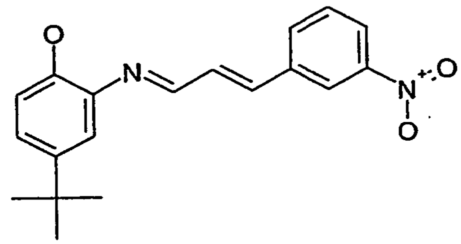
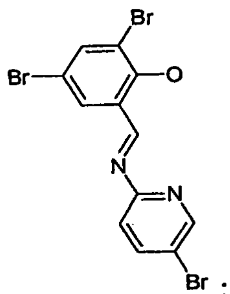
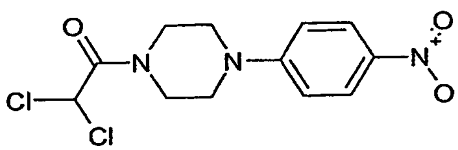
In yet other embodiments, the application provides methods of identifying an agent that inhibits the progression of a neurological disorder, comprising identifying

a test agent that disrupts a complex comprising a POSH polypeptide and a POSH-AP evaluating the effect of the test agent on the trafficking of a protein through the secretory pathway wherein an agent that disrupts localization of a POSH-AP is an agent that inhibits progression of a neurological disorder. In certain aspects the POSH-AP is HERPUD1.

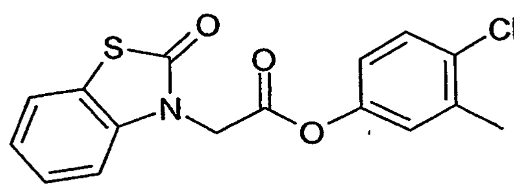
In yet other embodiments, this application provides methods of treating a viral infection in a subject in need thereof, comprising administering an agent that inhibits a POSH-AP in an amount sufficient to inhibit the viral infection. The agent is one that: inhibits a kinase activity of the POSH-AP; inhibits expression of the POSH-AP; inhibits the ubiquitin ligase activity of the POSH-AP; inhibits the phosphatase activity of the POSH-AP; inhibits the GTPase activity of the POSH-AP; and inhibits the ubiquitination of the POSH-AP. In certain embodiments, the POSH-AP comprises a polypeptide selected from the group consisting of: PKA, SNX1, SNX3, SMN1, SMN2, PTPN12, GOSR2, CENTB1, ARF1, ARF5, PACS-1, EPS8L2, HERPUD1, UNC84B, MSTP028, GOCAP, CBL-B, SYNE1, UBE2N (UBC13), SIAH1, TTC3, WASF1, HIP55, RALA, and SPG20. In certain aspects, the agent may be an siRNA construct, a small molecule, an antibody, or an antisense construct.

In certain embodiments, the agent is an siRNA construct comprising a nucleic acid sequence that hybridizes to an mRNA encoding the POSH-AP. Examples include siRNA constructs that inhibit the expression of HERPUD1 or MSTP028. Examples of siRNA constructs that inhibit the expression of HERPUD1 include: 5'GGAAGUUCUUCGGAACCUdTdT-3' and 5'- dTdTCCCUUCAAGAAGCCUUGGA-5'. Examples of siRNA constructs that inhibit the expression of MSTP028 include: 5'-AAGTGCTCACCGACAGTGAAG-3' and 5'-AAGATACTTATGAGCCTTTCT-3'.

In other aspects, the agents may be a small molecule inhibitor is selected from among the following categories: adenosine cyclic monophosphorothioate, isoquinolinesulfonamide, piperazine, piceatannol, and ellagic acid. In alternative embodiments, the agents may be a small molecule inhibitor that inhibits the ligase activity of a POSH polypeptide or inhinbits the ubiquitination of a POSH-AP. Examples of such small molecules include, for example:


;

;

;

;

;

;

and


.

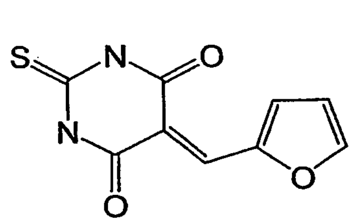
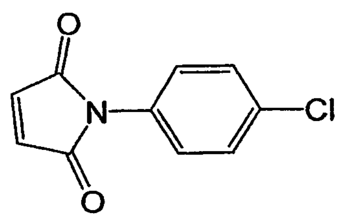
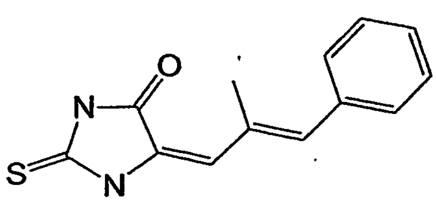
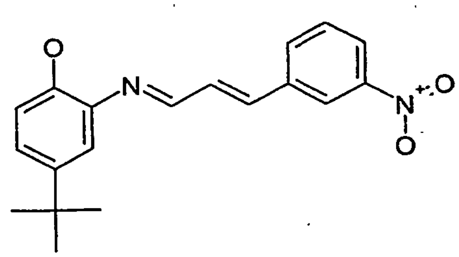
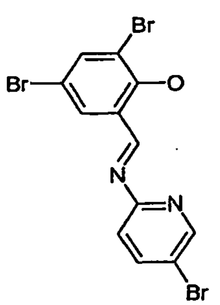
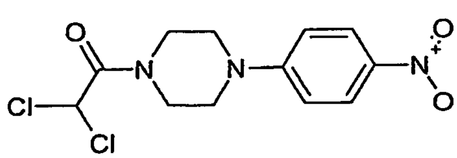
In certain embodiments, the application provides packaged pharmaceuticals for treating viral infections, comprising: a pharmaceutical composition comprising an inhibitor of a POSH-AP and a pharmaceutically acceptable carrier and instructions for use.

In certain embodiments, the application provides methods of treating or preventing a POSH associated cancer in a subject comprising administering an agent that inhibits a POSH-AP to a subject in need thereof, wherein said agent treats or

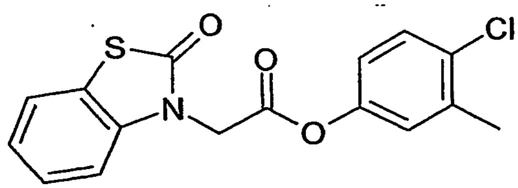
prevents cancer. The POSH-AP comprises a polypeptide selected from the group consisting of: PKA, SNX1, PTPN12, PPP1CA, CENTB1, ARF1, ARF5, EPS8L2, EIF3S3, CBL-B, RALA, SIAH1, TTC3, ATP6V0C, and VCY2IP1.

In yet other aspects, the application provides methods of treating a neurological disorder comprising administering an agent to a subject in need thereof, wherein said agent either inhibits the Ubiquitin ligase activity of POSH or inhibits the ubquitination of a POSH-AP. Examples of the POSH-AP include: PTPN12, DDEF1, EPS8L2, HERPUD1, GOCAP, CBL-B, SIAH1, SMN1, SMN2, TTC3, SPG20, SNX1, and ARF1.

Examples of the neurological disorders include Alzheimer's disease, Parkinson's disease, Huntington's disease, schizophrenia, Niemann-Pick's disease, and prion-associated diseases. In certain aspects, the agent is selected from the group consisting of: an siRNA construct, a small molecule, an antibody, and an antisense construct. Examples of the small molecules include:

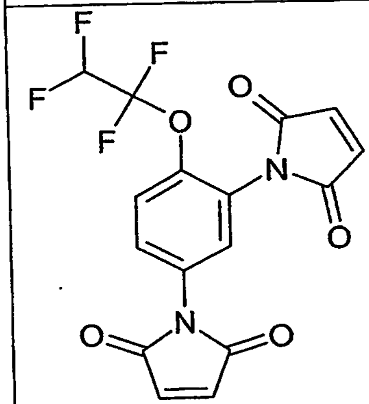
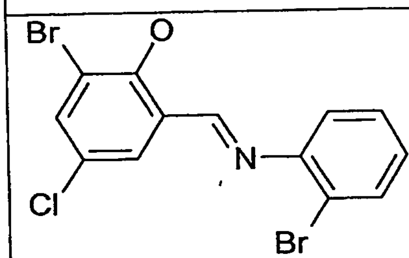

;

;

;

;

;

;

and

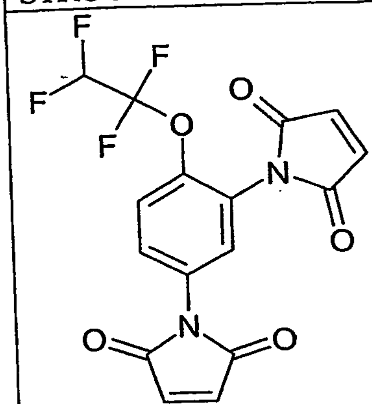
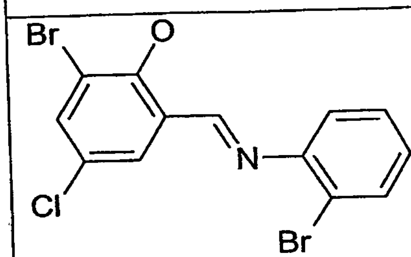


In certain aspects, the disclosure provides methods of treating viral hepatitis in a subject in need thereof. Such a method may comprise administering an effective amount of an agent that inhibits POSH or disrupts an interaction between POSH and a dynamin, preferably dynamin II. In certain embodiments, the subject has a viral hepatitis caused by HBV or HCV.

In certain aspects, the disclosure provides methods of inhibiting a hepatotrophic virus or a method for treating a disease associated with a hepatrophic virus, comprising administering an effective amount of an agent, wherein said agent inhibits POSH or an interaction between POSH and dynamin. In certain embodiments, the hepatrophic virus is selected from the group consisting of HAV, HBV, HCV, HDV, and HEV. The hepatotrophic virus associated disease may be, for example, viral hepatitis or hepatocellular carcinoma. An agent for any of the above methods may include, for example, a nucleic acid agent that decreases the level of POSH in cells of the subject (e.g., an antisense oligonucleotide, an RNAi construct, a DNA enzyme, a ribozyme) or small molecule inhibitors of POSH, as well as antibodies or other binding agents that bind to a surface of POSH or dynamin that participates in a POSH-dynamin interaction. An agent may be any of the following: a small molecule, an antibody, a fragment of an antibody, a peptidomimetic, and a polypeptide. Examples of small molecules include:

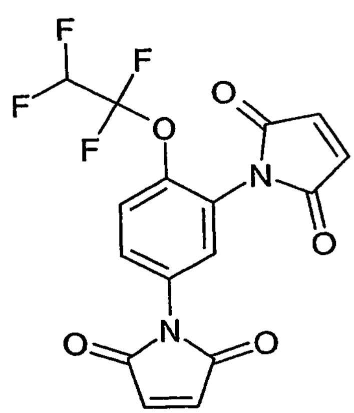
STRUCTURE	MW	CAS number
	384.2	14567-55-4
	389.5	414908-38-0

In certain embodiments, the application provides methods for inhibiting an HBV infection in a subject in need thereof, comprising administering an effective amount of a POSH inhibitor, wherein the HBV infection is inhibited in the subject. In additional embodiments, the disclosure provides methods for treating an HBV infection in a patient, comprising administering an effective amount of an agent that inhibits POSH or decreases the level of POSH protein or nucleic acid in an infected cell. An agent may be, for example, an RNAi construct that inhibits the expression of POSH. Optionally the RNAi construct is 20-25 nucleotides in length and optionally it is selected from any one of SEQ ID NOS: 15, 16, 18, 19, 21, 22, 24, and 25. The RNAi may be formulated as a liposome. An agent may be a small molecule inhibitor of POSH ubiquitin ligase activity, as disclosed herein. Examples of small molecule inhibitors of POSH include:

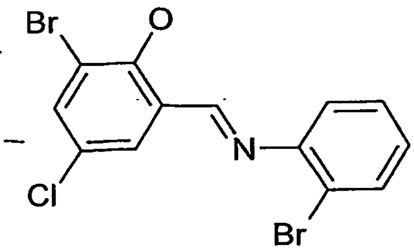
STRUCTURE	MW	CAS number
	384.2	14567-55-4
	389.5	414908-38-0

In certain aspects, the disclosure provides a method for treating an HBV infection in a patient, comprising administering an effective amount of an antisense oligonucleotide sufficient to bind a nucleic acid molecule, which nucleic acid molecule encodes a POSH polypeptide.

In certain embodiments, the application provides methods for inhibiting an HBV infection by administering an effective amount of a compound of the formula:



In additional embodiments, the application provides methods for treating an HBV infection by administering an effective amount of a compound of the formula:



In certain aspects, the disclosure provides methods for inhibiting the maturation of a lentivirus by modulating the activity of a Vpu polypeptide. In preferred embodiments, maturation of the lentivirus is inhibited by inhibiting the transport and/or assembly of viral particles in the TGN and from the TGN to the plasma membrane. A preferred lentivirus for application of such a method is the human immunodeficiency virus.

In certain aspects, the disclosure provides methods of inhibiting viral infection comprising administering an agent to a subject in need thereof, wherein said agent inhibits the interaction between a POSH polypeptide and Vpu.

In certain aspects, the disclosure provides methods for identifying a target polypeptide for antiviral therapy, the method comprising: a) selecting a test polypeptide known to localize or predicted to localize to the trans Golgi network; b) inhibiting an activity of the test polypeptide in a cell infected with a viral construct under conditions where, but for the inhibition of the activity of the test polypeptide, viral particles are released from the cell; and c) determining whether viral particles are released from the cell, wherein, if inhibiting the activity of the test polypeptide in the cell inhibits the release of viral particles from the cell, the test polypeptide is a target polypeptide for antiviral therapy. In a preferred embodiment, the test polypeptide is Vpu. Vpu activity may be inhibited, for example, by siRNA, antisense or other nucleic acid based method.

In certain aspects, the disclosure provides isolated, purified or recombinant complexes comprising a POSH polypeptide and a Vpu polypeptide. The POSH polypeptide may comprise, for example, a POSH SH3 domain, or a polypeptide at least 80% identical to such an SH3 domain. An antiviral agent may be selected based on its ability to disrupt a POSH-Vpu complex.

The practice of the present application will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology,

transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning*, Volumes I and II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis et al. U.S. Patent No: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu et al. eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

Other features and advantages of the application will be apparent from the following detailed description, and from the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows human POSH coding sequence (SEQ ID NO:1).

Figure 2 shows human POSH amino acid sequence (SEQ ID NO:2).

Figure 3 shows human POSH cDNA sequence (SEQ ID NO:3).

Figure 4 shows 5' cDNA fragment of human POSH (public gi:10432611; SEQ ID NO:4).

Figure 5 shows N terminus protein fragment of hPOSH (public gi:10432612; SEQ ID NO:5).

Figure 6 shows 3' mRNA fragment of hPOSH (public gi:7959248; SEQ ID NO:6).

Figure 7 shows C terminus protein fragment of hPOSH (public gi:7959249; SEQ ID NO:7).

Figure 8 shows human POSH full mRNA, annotated sequence.

Figure 9 shows domain analysis of human POSH.

Figure 10 is a diagram of human POSH nucleic acids. The diagram shows the full-length POSH gene and the position of regions amplified by RT-PCR or targeted by siRNA used in figure 11.

Figure 11 shows effect of knockdown of POSH mRNA by siRNA duplexes. HeLa SS-6 cells were transfected with siRNA against Lamin A/C (lanes 1, 2) or POSH (lanes 3-10). POSH siRNA was directed against the coding region (153 - lanes 3, 4; 155 - lanes 5, 6) or the 3'UTR (157 - lanes 7, 8; 159 - lanes 9, 10). Cells were harvested 24 hours post-transfection, RNA extracted, and POSH mRNA levels compared by RT-PCR of a discrete sequence in the coding region of the POSH gene (see figure 10). GAPDH is used an RT-PCR control in each reaction.

Figure 12 shows that POSH affects the release of VLP from cells. A) Phosphohimages of SDS-PAGE gels of immunoprecipitations of ^{35}S pulse-chase labeled Gag proteins are presented for cell and viral lysates from transfected HeLa cells that were either untreated or treated with POSH RNAi (50 nM for 48 hours). The time during the chase period (1, 2, 3, 4, and 5 hours after the pulse) are presented from left to right for each image.

Figure 13 shows release of VLP from cells at steady state. Hela cells were transfected with an HIV-encoding plasmid and siRNA. Lanes 1, 3 and 4 were transfected with wild-type HIV-encoding plasmid. Lane 2 was transfected with an HIV-encoding plasmid which contains a point mutation in p6 (PTAP to ATAP). Control siRNA (lamin A/C) was transfected to cells in lanes 1 and 2. siRNA to Tsg101 was transfected in lane 4 and siRNA to POSH in lane 3.

Figure 14 shows mouse POSH mRNA sequence (public gi:10946921; SEQ ID NO: 8).

Figure 15 shows mouse POSH Protein sequence (Public gi:10946922; SEQ ID NO: 9).

Figure 16 shows Drosophila melanogaster POSH mRNA sequence (public gi:17737480; SEQ ID NO:10).

Figure 17 shows Drosophila melanogaster POSH protein sequence (public gi:17737481; SEQ ID NO:11).

Figure 18 shows POSH domain analysis.

Figure 19 shows that human POSH has ubiquitin ligase activity.

Figure 20 shows that human POSH co-immunoprecipitates with RAC1.

Figure 21 shows that POSH knockdown results in decreased secretion of phospholipase D ("PLD").

Figure 22 shows effect of hPOSH on Gag-EGFP intracellular distribution.

Figure 23 shows intracellular distribution of HIV-1 Nef in hPOSH-depleted cells.

Figure 24 shows intracellular distribution of Src in hPOSH-depleted cells.

Figure 25 shows intracellular distribution of Rapsyn in hPOSH-depleted cells.

Figure 26 shows that POSH reduction by siRNA abrogates West Nile virus infectivity.

Figure 27 shows that POSH knockdown decreases the release of extracellular MMuLV particles.

Figure 28 shows that knock-down of human POSH entraps HIV virus particles in intracellular vesicles. HIV virus release was analyzed by electron microscopy following siRNA and full-length HIV plasmid transfection. Mature viruses were secreted by cells transfected with HIV plasmid and non-relevant siRNA (control, bottom panel). Knockdown of Tsg101 protein resulted in a budding defect, the viruses that were released had an immature phenotype (top panel). Knockdown of hPOSH levels resulted in accumulation of viruses inside the cell in intracellular vesicles (middle panel).

Figure 29A shows siRNA-mediated reduction of MSTP028 expression inhibits HIV virus-like particle production (Experiment 1).

Figure 29B shows siRNA-mediated reduction of MSTP028 expression inhibits HIV virus-like particle production (Experiment 2).

Figure 30 shows putative PKA phosphorylation sites in hPOSH. Amino acid sequence of hPOSH (70 residues per line). Motifs of the low stringency RxxS/T type are underlined. The high stringency motif R/KR/KxS/T is bordered. Putative S/T phosphorylation sites are highlighted in green. Color-coding of domains: Red – RING, Blue – SH3, Green – putative Rac-1 Binding Domain.

Figure 31 shows that phosphorylation of hPOSH regulates binding of GTP-loaded Rac-1. Bacterially expressed hPOSH (1 μg) (POSH) or GST (1 μg) (NS) were phosphorylated. Subsequently, GTPγS loaded or unloaded recombinant Rac-1 (0.2 μg) was added to hPOSH or GST. Bound rac1 was isolated as described in materials and methods and samples separated by SDS-PAGE on a 12% gel and immunobloted with anti-Rac-1. Input is 0.25 μg of Rac-1.

Figure 32 shows domain analysis of various POSH-APs.

Figure 33 shows siRNA-mediated reduction in HERPUD1 expression reduces HIV maturation.

Figure 34A shows that endogenous Herp levels are reduced in H153 cells. H153 (POSH-RNAi) and H187 (control RNAi) cells were transfected with a plasmid encoding Flag-ubiquitin. Total cell lysates (A) or Flag-immunoprecipitated material (B) were separated on 10% SDS-PAGE and immunoblotted with anti-Herp antibodies.

Figure 34B shows that exogenous Herp levels and its ubiquitination are reduced in POSH-depleted cells. H153 and H187 cells were co-transfected with Herp or control plasmids and a plasmid encoding Flag-ubiquitin (indicated above the figure). Total (A) and flag-immunoprecipitated material (B) were separated on 10% SDS-PAGE and immunoblotted with anti-Herp antibodies.

Figure 35 shows that the compounds CAS number 14567-55-4 and CAS number 414908-38-0 (lanes 7 and 8) inhibit HBV production.

Figure 36 provides the nucleic acid and amino acid sequences of POSH-APs.

DETAILED DESCRIPTION OF THE APPLICATION

1. Definitions

The term "binding" refers to a direct association between two molecules, due to, for example, covalent, electrostatic, hydrophobic, ionic and/or hydrogen-bond interactions under physiological conditions.

A "chimeric protein" or "fusion protein" is a fusion of a first amino acid sequence encoding a polypeptide with a second amino acid sequence defining a domain foreign to and not substantially homologous with any domain of the first amino acid sequence. A chimeric protein may present a foreign domain which is

found (albeit in a different protein) in an organism which also expresses the first protein, or it may be an "interspecies", "intergenic", etc. fusion of protein structures expressed by different kinds of organisms.

The terms "compound", "test compound" and "molecule" are used herein interchangeably and are meant to include, but are not limited to, peptides, nucleic acids, carbohydrates, small organic molecules, natural product extract libraries, and any other molecules (including, but not limited to, chemicals, metals and organometallic compounds).

The phrase "conservative amino acid substitution" refers to grouping of amino acids on the basis of certain common properties. A functional way to define common properties between individual amino acids is to analyze the normalized frequencies of amino acid changes between corresponding proteins of homologous organisms (Schulz, G. E. and R. H. Schirmer., Principles of Protein Structure, Springer-Verlag). According to such analyses, groups of amino acids may be defined where amino acids within a group exchange preferentially with each other, and therefore resemble each other most in their impact on the overall protein structure (Schulz, G. E. and R. H. Schirmer, Principles of Protein Structure, Springer-Verlag). Examples of amino acid groups defined in this manner include:

(i) a charged group, consisting of Glu and Asp, Lys, Arg and His,

(ii) a positively-charged group, consisting of Lys, Arg and His,

(iii) a negatively-charged group, consisting of Glu and Asp,

(iv) an aromatic group, consisting of Phe, Tyr and Trp,

(v) a nitrogen ring group, consisting of His and Trp,

(vi) a large aliphatic nonpolar group, consisting of Val, Leu and Ile,

(vii) a slightly-polar group, consisting of Met and Cys,

(viii) a small-residue group, consisting of Ser, Thr, Asp, Asn, Gly, Ala, Glu, Gln and Pro,

(ix) an aliphatic group consisting of Val, Leu, Ile, Met and Cys, and

(x) a small hydroxyl group consisting of Ser and Thr.

In addition to the groups presented above, each amino acid residue may form its own group, and the group formed by an individual amino acid may be referred to

simply by the one and/or three letter abbreviation for that amino acid commonly used in the art.

A "conserved residue" is an amino acid that is relatively invariant across a range of similar proteins. Often conserved residues will vary only by being replaced with a similar amino acid, as described above for "conservative amino acid substitution".

The term "domain" as used herein refers to a region of a protein that comprises a particular structure and/or performs a particular function.

The term "envelope virus" as used herein refers to any virus that uses cellular membrane and/or any organelle membrane in the viral release process.

"Homology" or "identity" or "similarity" refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology and identity can each be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When an equivalent position in the compared sequences is occupied by the same base or amino acid, then the molecules are identical at that position; when the equivalent site occupied by the same or a similar amino acid residue (e.g., similar in steric and/or electronic nature), then the molecules can be referred to as homologous (similar) at that position. Expression as a percentage of homology/similarity or identity refers to a function of the number of identical or similar amino acids at positions shared by the compared sequences. A sequence which is "unrelated" or "non-homologous" shares less than 40% identity, though preferably less than 25% identity with a sequence of the present application. In comparing two sequences, the absence of residues (amino acids or nucleic acids) or presence of extra residues also decreases the identity and homology/similarity.

The term "homology" describes a mathematically based comparison of sequence similarities which is used to identify genes or proteins with similar functions or motifs. The nucleic acid and protein sequences of the present application may be used as a "query sequence" to perform a search against public databases to, for example, identify other family members, related sequences or homologs. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) J Mol. Biol. 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score=100,

wordlength=12 to obtain nucleotide sequences homologous to nucleic acid molecules of the application. BLAST protein searches can be performed with the XBLAST program, score=50, wordlength=3 to obtain amino acid sequences homologous to protein molecules of the application. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) Nucleic Acids Res. 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and BLAST) can be used. See http://www.ncbi.nlm.nih.gov.

As used herein, "identity" means the percentage of identical nucleotide or amino acid residues at corresponding positions in two or more sequences when the sequences are aligned to maximize sequence matching, i.e., taking into account gaps and insertions. Identity can be readily calculated by known methods, including but not limited to those described in (Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., Nucleic Acids Research 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Altschul, S. F. et al., J. Molec. Biol. 215: 403-410 (1990) and Altschul et al. Nuc. Acids Res. 25: 3389-3402 (1997)). The BLAST X program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, Md. 20894; Altschul, S., et al., J. Mol. Biol. 215: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.

The term "isolated", as used herein with reference to the subject proteins and protein complexes, refers to a preparation of protein or protein complex that is

essentially free from contaminating proteins that normally would be present with the protein or complex, e.g., in the cellular milieu in which the protein or complex is found endogenously. Thus, an isolated protein complex is isolated from cellular components that normally would "contaminate" or interfere with the study of the complex in isolation, for instance while screening for modulators thereof. It is to be understood, however, that such an "isolated" complex may incorporate other proteins the modulation of which, by the subject protein or protein complex, is being investigated.

The term "isolated" as also used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules in a form which does not occur in nature. Moreover, an "isolated nucleic acid" is meant to include nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state.

Lentiviruses include primate lentiviruses, e.g., human immunodeficiency virus types 1 and 2 (HIV-1/HIV-2); simian immunodeficiency virus (SIV) from Chimpanzee (SIVcpz), Sooty mangabey (SIVsmm), African Green Monkey (SIVagm), Syke's monkey (SIVsyk), Mandrill (SIVmnd) and Macaque (SIVmac). Lentiviruses also include feline lentiviruses, e.g., Feline immunodeficiency virus (FIV); Bovine lentiviruses, e.g., Bovine immunodeficiency virus (BIV); Ovine lentiviruses, e.g., Maedi/Visna virus (MVV) and Caprine arthritis encephalitis virus (CAEV); and Equine lentiviruses, e.g., Equine infectious anemia virus (EIAV). All lentiviruses express at least two additional regulatory proteins (Tat, Rev) in addition to Gag, Pol, and Env proteins. Primate lentiviruses produce other accessory proteins including Nef, Vpr, Vpu, Vpx, and Vif. Generally, lentiviruses are the causative agents of a variety of disease, including, in addition to immunodeficiency, neurological degeneration, and arthritis. Nucleotide sequences of the various lentiviruses can be found in Genbank under the following Accession Nos. (from J. M. Coffin, S. H. Hughes, and H. E. Varmus, "Retroviruses" Cold Spring Harbor Laboratory Press, 1997, p 804): 1) HIV-1: K03455, M19921, K02013, M38431, M38429, K02007 and M17449; 2) HIV-2: M30502, J04542, M30895, J04498, M15390, M31113 and L07625; 3) SIV:M29975, M30931, M58410, M66437, L06042, M33262, M19499, M32741, M31345 and L03295; 4) FIV: M25381,

M36968 and Ul 1820; 5)BIV. M32690; 6) E1AV: M16575, M87581 and U01866; 6) Visna: M10608, M51543, L06906, M60609 and M60610; 7) CAEV: M33677; and 8) Ovine lentivirus M31646 and M34193. Lentiviral DNA can also be obtained from the American Type Culture Collection (ATCC). For example, feline immunodeficiency virus is available under ATCC Designation No. VR-2333 and VR-3112. Equine infectious anemia virus A is available under ATCC Designation No. VR-778. Caprine arthritis-encephalitis virus is available under ATCC Designation No. VR-905. Visna virus is available under ATCC Designation No. VR-779.

As used herein, the term "nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single-stranded (such as sense or antisense) and double-stranded polynucleotides.

The term "maturation" as used herein refers to the production, post-translational processing, assembly and/or release of proteins that form a viral particle. Accrodingly, this includes the processing of viral proteins leading to the pinching off of nascent virion from the cell membrane.

A "POSH nucleic acid" is a nucleic acid comprising a sequence as represented in any of SEQ ID Nos:1, 3, 4, 6, 8, and 10 as well as any of the variants described herein.

A "POSH polypeptide" or "POSH protein" is a polypeptide comprising a sequence as represented in any of SEQ ID Nos: 2, 5, 7, 9 and11 as well as any of the variations described herein.

A "POSH-associated protein" or "POSH-AP" refers to a protein capable of interacting with and/or binding to a POSH polypeptide. Generally, the POSH-AP may interact directly or indirectly with the POSH polypeptide. Preferred POSH-APs include those provided in Table 7. Other preferred POSH-APs include those listed in Table 8. Examples of these and other POSH-APs are provided throughout.

The terms peptides, proteins and polypeptides are used interchangeably herein.

The term "purified protein" refers to a preparation of a protein or proteins which are preferably isolated from, or otherwise substantially free of, other proteins normally associated with the protein(s) in a cell or cell lysate. The term "substantially free of other cellular proteins" (also referred to herein as "substantially free of other contaminating proteins") is defined as encompassing individual preparations of each of the component proteins comprising less than 20% (by dry weight) contaminating protein, and preferably comprises less than 5% contaminating protein. Functional forms of each of the component proteins can be prepared as purified preparations by using a cloned gene as described in the attached examples. By "purified", it is meant, when referring to component protein preparations used to generate a reconstituted protein mixture, that the indicated molecule is present in the substantial absence of other biological macromolecules, such as other proteins (particularly other proteins which may substantially mask, diminish, confuse or alter the characteristics of the component proteins either as purified preparations or in their function in the subject reconstituted mixture). The term "purified" as used herein preferably means at least 80% by dry weight, more preferably in the range of 85% by weight, more preferably 95-99% by weight, and most preferably at least 99.8% by weight, of biological macromolecules of the same type present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 5000, can be present). The term "pure" as used herein preferably has the same numerical limits as "purified" immediately above.

A "recombinant nucleic acid" is any nucleic acid that has been placed adjacent to another nucleic acid by recombinant DNA techniques. A "recombined nucleic acid" also includes any nucleic acid that has been placed next to a second nucleic acid by a laboratory genetic technique such as, for example, tranformation and integration, transposon hopping or viral insertion. In general, a recombined nucleic acid is not naturally located adjacent to the second nucleic acid.

The term "recombinant protein" refers to a protein of the present application which is produced by recombinant DNA techniques, wherein generally DNA encoding the expressed protein is inserted into a suitable expression vector which is

in turn used to transform a host cell to produce the heterologous protein. Moreover, the phrase "derived from", with respect to a recombinant gene encoding the recombinant protein is meant to include within the meaning of "recombinant protein" those proteins having an amino acid sequence of a native protein, or an amino acid sequence similar thereto which is generated by mutations including substitutions and deletions of a naturally occurring protein.

A "RING domain" or "Ring Finger" is a zinc-binding domain with a defined octet of cysteine and histidine residues. Certain RING domains comprise the consensus sequences as set forth below (amino acid nomenclature is as set forth in Table 1): Cys Xaa Xaa Cys Xaa_{10-20} Cys Xaa His Xaa_{2-5} Cys Xaa Xaa Cys Xaa_{13-50} Cys Xaa Xaa Cys or Cys Xaa Xaa Cys Xaa_{10-20} Cys Xaa His Xaa_{2-5} His Xaa Xaa Cys Xaa_{13-50} Cys Xaa Xaa Cys. Certain RING domains are represented as amino acid sequences that are at least 80% identical to amino acids 12-52 of SEQ ID NO: 2 and is set forth in SEQ ID No: 26. Preferred RING domains are 85%, 90%, 95%, 98% and, most preferably, 100% identical to the amino acid sequence of SEQ ID NO: 26. Preferred RING domains of the application bind to various protein partners to form a complex that has ubiquitin ligase activity. RING domains preferably interact with at least one of the following protein types: F box proteins, E2 ubiquitin conjugating enzymes and cullins.

The term "RNA interference" or "RNAi" refers to any method by which expression of a gene or gene product is decreased by introducing into a target cell one or more double-stranded RNAs which are homologous to the gene of interest (particularly to the messenger RNA of the gene of interest). RNAi may also be achieved by introduction of a DNA:RNA hybrid wherein the antisense strand (relative to the target) is RNA. Either strand may include one or more modifications to the base or sugar-phosphate backbone. Any nucleic acid preparation designed to achieve an RNA interference effect is referred to herein as an siRNA construct. Phosphorothioate is a particularly common modification to the backbone of an siRNA construct.

"Small molecule" as used herein, is meant to refer to a composition, which has a molecular weight of less than about 5 kD and most preferably less than about 2.5 kD. Small molecules can be nucleic acids, peptides, polypeptides,

peptidomimetics, carbohydrates, lipids or other organic (carbon containing) or inorganic molecules. Many pharmaceutical companies have extensive libraries of chemical and/or biological mixtures comprising arrays of small molecules, often fungal, bacterial, or algal extracts, which can be screened with any of the assays of the application.

An "SH3" or "Src Homology 3" domain is a protein domain of generally about 60 amino acid residues first identified as a conserved sequence in the non-catalytic part of several cytoplasmic protein tyrosine kinases (e.g., Src, Abl, Lck). SH3 domains mediate assembly of specific protein complexes via binding to proline-rich peptides. Exemplary SH3 domains are represented by amino acids 137-192, 199-258, 448-505 and 832-888 of SEQ ID NO:2 and are set forth in SEQ ID Nos: 27-30. In certain embodiments, an SH3 domain interacts with a consensus sequence of RXaaXaaPXaaX6P (where X6, as defined in table 1 below, is a hydrophobic amino acid). In certain embodiments, an SH3 domain interacts with one or more of the following sequences: P(T/S)AP, PFRDY, RPEPTAP, RQGPKEP, RQGPKEPFR, RPEPTAPEE and RPLPVAP.

As used herein, the term "specifically hybridizes" refers to the ability of a nucleic acid probe/primer of the application to hybridize to at least 12, 15, 20, 25, 30, 35, 40, 45, 50 or 100 consecutive nucleotides of a POSH sequence, or a sequence complementary thereto, or naturally occurring mutants thereof, such that it has less than 15%, preferably less than 10%, and more preferably less than 5% background hybridization to a cellular nucleic acid (e.g., mRNA or genomic DNA) other than the POSH gene. A variety of hybridization conditions may be used to detect specific hybridization, and the stringency is determined primarily by the wash stage of the hybridization assay. Generally high temperatures and low salt concentrations give high stringency, while low temperatures and high salt concentrations give low stringency. Low stringency hybridization is achieved by washing in, for example, about 2.0 x SSC at 50 °C, and high stringency is acheived with about 0.2 x SSC at 50 °C. Further descriptions of stringency are provided below.

As applied to polypeptides, "substantial sequence identity" means that two peptide sequences, when optimally aligned, such as by the programs GAP or

BESTFIT using default gap which share at least 90 percent sequence identity, preferably at least 95 percent sequence identity, more preferably at least 99 percent sequence identity or more. Preferably, residue positions which are not identical differ by conservative amino acid substitutions. For example, the substitution of amino acids having similar chemical properties such as charge or polarity are not likely to effect the properties of a protein. Examples include glutamine for asparagine or glutamic acid for aspartic acid.

As is well known, genes for a particular polypeptide may exist in single or multiple copies within the genome of an individual. Such duplicate genes may be identical or may have certain modifications, including nucleotide substitutions, additions or deletions, which all still code for polypeptides having substantially the same activity.

A "virion" is a complete viral particle; nucleic acid and capsid (and a lipid envelope in some viruses. A "viral particle" may be incomplete, as when produced by a cell transfected with a defective virus (e.g., an HIV virus-like particle system).

Table 1: Abbreviations for classes of amino acids*

Symbol	**Category**	**Amino Acids Represented**
X1	Alcohol	Ser, Thr
X2	Aliphatic	Ile, Leu, Val
Xaa	Any	Ala, Cys, Asp, Glu, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, Tyr
X4	Aromatic	Phe, His, Trp, Tyr
X5	Charged	Asp, Glu, His, Lys, Arg

X6	Hydrophobic	Ala, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Thr, Val, Trp, Tyr
X7	Negative	Asp, Glu
X8	Polar	Cys, Asp, Glu, His, Lys, Asn, Gln, Arg, Ser, Thr
X9	Positive	His, Lys, Arg
X10	Small	Ala, Cys, Asp, Gly, Asn, Pro, Ser, Thr, Val
X11	Tiny	Ala, Gly, Ser
X12	Turnlike	Ala, Cys, Asp, Glu, Gly, His, Lys, Asn, Gln, Arg, Ser, Thr
X13	Asparagine-Aspartate	Asn, Asp

* Abbreviations as adopted from http://smart.embl-heidelberg.de/SMART_DATA/alignments/consensus/grouping.html.

2. Overview

In certain aspects, the application relates to the discovery of novel associations between POSH proteins and other proteins (termed POSH-APs), and related methods and compositions. In certain aspects, the application relates to novel associations among certain disease states, POSH nucleic acids and proteins, and POSH-AP nucleic acids and proteins.

In certain aspects, by identifying proteins associated with POSH, and particularly human POSH, the present application provides a conceptual link between the POSH-APs and cellular processes and disorders associated with POSH-APs, and POSH itself. Accordingly, in certain embodiments of the disclosure, agents that modulate a POSH-AP may now be used to modulate POSH functions

and disorders associated with POSH function, such as viral disorders, POSH-associated cancers, and POSH-associated neural disorders. Additionally, test agents may be screened for an effect on a POSH-AP and then further tested for an effect on a POSH function or a disorder associated with POSH function. Likewise, in certain embodiments of the disclosure, agents that modulate POSH may now be used to modulate POSH-AP functions and disorders associated with POSH-AP function, including a variety of cancers. Additionally, test agents may be screened for an effect on POSH and then further tested for effect on a POSH-AP function or a disorder associated with POSH-AP function. In further aspects, the application provides nucleic acid agents (e.g., RNAi probes, antisense nucleic acids), antibody-related agents, small molecules and other agents that affect POSH function, and the use of same in modulating POSH and/or POSH-AP activity.

POSH intersects with and regulates a wide range of key cellular functions that may be manipulated by affecting the level of and/or activity of POSH polypeptides or POSH-AP polypeptides. Many features of POSH, and particularly human POSH, are described in PCT patent publications WO03/095971A2 (application no. WO2002US0036366) and WO03/078601A2 (application no. WO2003US0008194) the teachings of which are incorporated by reference herein.

As described in the above-referenced publications, native human POSH is a large polypeptide containing a RING domain and four SH3 domains. POSH is a ubiquitin ligase (also termed an "E3" enzyme); the RING domain mediates ubiquitination of, for example, the POSH polypeptide itself. POSH interacts with a large number of proteins and participates in a host of different biological processes. As demonstrated in this disclosure, POSH associates with a number of different proteins in the cell. POSH co-localizes with proteins that are known to be located in the trans-Golgi network, implying that POSH participates in the trafficking of proteins in the secretory system. The term "secretory system" should be understood as referring to the membrane compartments and associated proteins and other molecules that are involved in the movement of proteins from the site of translation to a location within a vacuole, a compartment in the secretory pathway itself, a lysosome or endosome or to a location at the plasma membrane or outside the cell. Commonly cited examples of compartments in the secretory system include the

endoplasmic reticulum, the Golgi apparatus and the cis and trans Golgi networks. In addition, Applicants have demonstrated that POSH is necessary for proper secretion, localization or processing of a variety of proteins, including phospholipase D, HIV Gag, HIV Nef, Rapsyn and Src. Many of these proteins are myristoylated, indicating that POSH plays a general role in the processing and proper localization of myristoylated proteins. N-myristoylation is an acylation process, which results in covalent attachment of myristate, a 14-carbon saturated fatty acid to the N-terminal glycine of proteins (Farazi et al., J. Biol. Chem. 276: 39501-04 (2001)). N-myristoylation occurs co-translationaly and promotes weak and reversible protein-membrane interaction. Myristoylated proteins are found both in the cytoplasm and associated with membrane. Membrane association is dependent on protein configuration, i.e., surface accessibility of the myristoyl group may be regulated by protein modifications, such as phosphorylation, ubiquitination etc. Modulation of intracellular transport of myristoylated proteins in the application includes effects on transport and localization of these modified proteins.

As described herein, POSH and POSH-APs are involved in viral maturation, including the production, post-translational processing, assembly and/or release of proteins in a viral particle. Accordingly, viral infections may be ameliorated by inhibiting an activity (e.g., ubiquitin ligase activity or target protein interaction) of POSH or a POSH-AP (e.g., inhibition of kinase activity or ubiquitin ligase activity), and in preferred embodiments, the virus is a retroid virus, an RNA virus or an envelope virus, including HIV, Ebola, HBV, HCV, HTLV, West Nile Virus (WNV) or Moloney Murine Leukemia Virus (MMuLV). Additional viral species are described in greater detail below. In certain instances, a decrease of a POSH function is lethal to cells infected with a virus that employs POSH in release of viral particles.

In certain aspects, the application describes an hPOSH interaction with Rac, a small GTPase and the POSH associated kinases MLK, MKK and JNK. Rho, Rac and Cdc42 operate together to regulate organization of the actin cytoskeleton and the MLK-MKK-JNK MAP kinase pathway (referred to herein as the "JNK pathway" or "Rac-JNK pathway" (Xu et al., 2003, EMBO J. 2: 252-61). Ectopic expression of mouse POSH ("mPOSH") activates the JNK pathway and causes nuclear

localization of NF-κB. Overexpression of mPOSH in fibroblasts stimulates apoptosis. (Tapon et al. (1998) EMBO J. 17:1395-404). In Drosophila, POSH may interact with, or otherwise influence the signaling of, another GTPase, Ras. (Schnorr et al. (2001) Genetics 159: 609-22). The JNK pathway and NF-κB regulate a variety of key genes involved in, for example, immune responses, inflammation, cell proliferation and apoptosis. For example, NF-κB regulates the production of interleukin 1, interleukin 8, tumor necrosis factor and many cell adhesion molecules. NF-κB has both pro-apoptotic and anti-apoptotic roles in the cell (e.g., in FAS-induced cell death and TNF-alpha signaling, respectively). NF-κB is negatively regulated, in part, by the inhibitor proteins IκBα and IκBβ (collectively termed "IκB"). Phosphorylation of IκB permits activation and nuclear localization of NF-κB. Phosphorylation of IκB triggers its degradation by the ubiquitin system. In an additional embodiment, a POSH polypeptide promotes nuclear localization of NF-κB. In further embodiments, manipulation of POSH levels and/or activities may be used to manipulate apoptosis. By upregulating POSH or a POSH-AP, apoptosis may be stimulated in certain cells, and this will generally be desirable in conditions characterized by excessive cell proliferation (e.g., in certain cancers). By downregulating POSH or a POSH-AP, apoptosis may be diminished in certain cells, and this will generally be desirable in conditions characterized by excessive cell death, such as myocardial infarction, stroke, degenerative diseases of muscle and nerve (particularly Alzheimer's disease), and for organ preservation prior to transplant. In a further embodiment, a POSH polypeptide associates with a vesicular trafficking complex, such as a clathrin- or coatomer- containing complex, and particularly a trafficking complex that localizes to the nucleus and/or Golgi apparatus.

As described in WO03/078601A2 (application no. WO2003US0008194), POSH is overexpressed in a variety of cancers, and downregulation of POSH is associated with a decrease in proliferation in at least one cancer cell line. Accordingly, agents that modulate POSH itself or a POSH-AP may be used to treat POSH associated cancers. POSH associated cancers include those cancers in which POSH is overexpressed and/or in which downregulation of POSH leads to a

decrease in the proliferation or survival of cancer cells. POSH-associated cancers are described in more detail below. In addition, it is notable that many proteins shown herein to be affected by POSH downregulation are themselves involved in cancers. Phospholipase D and SRC are both aberrantly processed in a POSH-impaired cell, and therefore modulation of POSH and/or a POSH-AP may affect the wide range of cancers in which PLD and SRC play a significant role.

As described in WO03/095971A2 (application no. WO2002US0036366) and WO03/078601A2 (application no. WO2003US0008194), POSH polypeptides function as E3 enzymes in the ubiquitination system. Accordingly, downregulation or upregulation of POSH ubiquitin ligase activity can be used to manipulate biological processes that are affected by protein ubiquitination. Modulation of POSH ubiquitin ligase activity may be used to affect POSH-APs and related biological processes, and likewise, modulation of POSH-APs may be used to affect POSH ubiquitin ligase activity and related processes. Downregulation or upregulation may be achieved at any stage of POSH formation and regulation, including transcriptional, translational or post-translational regulation. For example, POSH transcript levels may be decreased by RNAi targeted at a POSH gene sequence. As another example, POSH ubiquitin ligase activity may be inhibited by contacting POSH with an antibody that binds to and interferes with a POSH RING domain or a domain of POSH that mediates interaction with a target protein (a protein that is ubiquitinated at least in part because of POSH activity). As a further example, small molecule inhibitors of POSH ubiquitin ligase activity are provided herein. As another example, POSH activity may be increased by causing increased expression of POSH or an active portion thereof. POSH, and POSH-APs that modulate POSH ubiquitin ligase activity may participate in biological processes including, for example, one or more of the various stages of a viral lifecycle, such as viral entry into a cell, production of viral proteins, assembly of viral proteins and release of viral particles from the cell. POSH may participate in diseases characterized by the accumulation of ubiquitinated proteins, such as dementias (e.g., Alzheimer's and Pick's), inclusion body myositis and myopathies, polyglucosan body myopathy, and certain forms of amyotrophic lateral sclerosis. POSH may

participate in diseases characterized by excessive or inappropriate ubiquitination and/or protein degradation.

3. POSH Associated Proteins

In certain aspects, the application relates to the discovery of novel associations between POSH proteins and other proteins (termed POSH-APs), and related methods and compositions. In certain aspects, the application relates to novel associations among certain disease states, POSH nucleic acids and proteins, and POSH-AP nucleic acids and proteins. POSH-APs may interact either directly or indirectly with POSH. In certain embodiments, a POSH-AP binds directly to a POSH polypeptide.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with one subunit of Protein Kinase A (PKA; cAMP-dependent protein kinase). In one aspect, the application relates to the discovery that POSH binds directly with PRKAR1A. This interaction was identified by Applicants in a yeast 2-hybrid assay. Exemplary PKA subunits may include, but are not limited to, a regulatory subunit (e.g., PRKAR1A) and a catalytic subunit (e.g., PRKACA or PRKACB). PKA is an essential enzyme in the signaling pathway of the second messenger cyclic AMP (cAMP). Through phosphorylation of target proteins, PKA controls many biochemical events in the cell including regulation of metabolism, ion transport, and gene transcription. The PKA holoenzyme is composed of two regulatory and two catalytic subunits and dissociates from the regulatory subunits upon binding of cAMP. The PKA enzyme is inactive in the absence of cAMP. Activation of PKA occurs when two cAMP molecules bind to each regulatory subunit, eliciting a reversible conformational change that releases active catalytic subunits.

A number of human PKA subunits have been characterized, including a regulatory subunit (type I alpha: PRKAR1) and two catalytic subunits (C-alpha: PRKACA; and C-beta: PRKACB). Boshart et al. identified the regulatory subunit PRKAR1 of PKA as the product of the TSE1 locus (Boshart, M et al. (1991) Cell 66: 849-859). The evidence consisted of concordant expression of PRKAR1 mRNA and TSE1 genetic activity, high resolution physical mapping of the two genes on human chromosome 17, and the ability of transfected PRKAR1 cDNA to generate a

phenocopy of TSE1-mediated extinction. Jones et al. independently established identity of TSE1 and the RI-alpha subunit (Jones, KW et al. (1991) Cell 66: 861-872).

Other than a role of PKA in metabolism, PKA subunits have recently been implicated in multiple diseases. For example, a specific role for localized PRKAR1 has been demonstrated in human T lymphocytes, where type I PKA localizes to the activated TCR complex and is required for attenuation of signals propagated through this complex (Skalhegg, BS et al. (1992) J Biol Chem 267:15707-15714; Skalhegg, BS et al. (1994) Science 263: 84-87). The importance of type I PKA-mediated effects in attenuation of T cell replication has led to its consideration as a therapeutic target in combined variable immunodeficiency (CVI) and acquired immune deficiency syndrome (AIDS). Furthermore, type I PKA in T cells may also serve as a potential therapeutic target in systemic lupus erythematosis (SLE). For example, a series of recently published articles has uncovered the first human disease mapping to a PKA subunit-Carney complex (Casey, M et al. (2000) J Clin Invest 106: R31-38; Kirschner, LS et al. (2000) Nat Genet 26: 89-92). Carney complex (CNC) is a multiple neoplasia syndrome characterized by spotty skin pigmentation, cardiac and skin myxomas, endocrine tumors, and psammomatous melanotic schwannomas. CNC maps to two genomic loci, 17q24 and 2p16. Familial cases mapping to the 17q24 locus reveal deletions/mutations in the PRKAR1 coding exons leading to frameshifts and premature stop codons—no mRNA and protein from the mutant alleles has been observed.

Accordingly, in certain aspects of the present disclosure, POSH participates in the formation of PKA complexes, including human PKA-containing complexes. Certain POSH polypeptides may be involved in disorders of the immune system, e.g., autoimmune disorders. Certain POSH polypeptides may be involved in the regulation of T-cell activation. In certain aspects, POSH participates in the ubiquitination of PI3K. In certain aspects, PKA subunit polypeptides participate in POSH-mediated processes.

Additionally, the disclosure relates in part to the discovery that PKA phosphorylates POSH, and further, that this phosphorylation inhibits the interaction of POSH with small GTPases, such as Rac. Small GTPases are important in

vesicular trafficking, and therefore the findings disclosed herein demonstrate that POSH phosphorylation regulates the formation of complexes between POSH and proteins involved in the secretory system, such as Rac, TCL, TC10, Cdc42, Wrch-1, Rac2, Rac3 or RhoG. Applicants have shown that inhibition of PKA and POSH has similar effects, indicating that inhibition of PKA will achieve an effect similar to that of inhibition of POSH. However, given the effect of PKA on POSH interaction with proteins in the secretory pathway, it is expected that PKA regulates the timing of cyclical interactions that are needed to effect vesicular trafficking. Accordingly, it is expected that significant inhibition or activation of PKA will cause a disruption in POSH function.

The term "PKA subunit" is used herein to refer to a full-length human PKA subunit which includes a regulatory subunit (e.g., PRKAR1A) and a catalytic subunit (e.g., PRKACB or PRKACA), as well as an alternative PKA subunit composed of separate PKA subunit sequences (e.g., nucleic acid sequences) that may be a splice variant. The term "PKA subunit" is used herein to refer as well to various naturally occurring PKA subunit homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring PKA subunit (e.g., SEQ ID NOs: 264-265, 111-122, 395-396). The term specifically includes human PKA subunit nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with human UNC84B, a human homolog of C. elegans Unc-84. Accordingly, the application provides complexes comprising POSH and UNC84B. In one aspect, the application relates to the discovery that POSH binds directly with UNC84B. This interaction was identified by Applicants in a yeast 2-hybrid assay. In C. elegans, Unc-84 is involved in the cellular positioning of the nucleus. UNC84/SUN is positioned at the nuclear membrane and recruits Syne/ANC-1, which directly tethers the nuclear envelope to the actin cytoskeleton. Accordingly, in certain aspects, POSH participates in formation of a UNC84 complexes, including human UNC84B-containing complexes, and in the connections between the nucleus and the cytoskeleton. In certain aspects, UNC84

polypeptides participate in POSH-mediated processes. See, for example, Starr and Han, 2003, J Cell Sci 116(Pt 2):211-6.

The term UNC84 is used herein to refer to various naturally occurring Unc-84 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring UNC84 (e.g., SEQ ID NOs: 314, 211-213). The term specifically includes human UNC84B nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with human GOCAP1. Certain GOCAP1 polypeptides are cytoplasmic proteins associated with the Golgi complex. Accordingly, the application provides complexes comprising POSH and GOCAP1. In one aspect, the application relates to the discovery that POSH binds directly with GOCAP1. This interaction was identified by Applicants in a yeast 2-hybrid assay. In certain aspects, these complexes associate with the Golgi complex. GOCAP1 is synonymous with GCP60. Certain GCP60 polypeptides interact with the Golgi complex integral membrane protein, giantin. Certain GCP60 polypeptides are involved in the maintenance of the Golgi structure through interaction with giantin and affect protein transport between the endoplasmic reticulum and the Golgi complex (Sohda, M, et al. (2001) J Biol Chem 276:45298-306). In certain aspects, GOCAP1 polypeptides participate in POSH-mediated processes.

The term GOCAP1 is used herein to refer to various naturally occurring GOCAP1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring GOCAP1 (e.g., SEQ ID NOs: 240-243, 61-68). The term specifically includes human GOCAP1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with human PTPN12, a protein tyrosine phosphatase. Accordingly, the application provides complexes comprising POSH and PTPN12. In one aspect, the application relates to the discovery that POSH binds directly with PTPN12. This interaction was identified by Applicants in a yeast 2-hybrid assay.

PTPN12 polypeptides are synonymous with the protein tyrosine phosphatase, PTP-PEST. PTP-PEST polypeptides contain proline-rich sequences and are rich in proline, glutamate, serine, and threonine residues at their carboxyl terminus, features characteristic of PEST motifs. Certain PTP-PEST polypeptides interact with paxillin, a scaffolding protein to which focal adhesion proteins bind, leading to the formation of the focal adhesion contact (Shen, Y et al. (1998) J Biol Chem 273:6474-81). Certain PTP-PEST polypeptides associate with the focal adhesion protein, p130cas (Garton, AJ et al. (1997) Oncogene 15:877-85). Certain PTP-PEST polypeptides have also been shown to associate with JAK2, PSTPIP and WASP, gelsolin, cell adhesion kinase beta, Csk, Hef 1 or Sin , Hic-5, or Shc (See, for example, Horsch, et al (2001) Mol Endocrinol 15:2182-96; Cote, et al (2002) J Biol Chem 277:2973-86; Chellaiah, et al (2001) J Biol Chem 276:47434-44; Lyons, et al (2001) J Biol Chem 276:24422-31; Davidson, et al (1997) J Biol Chem 21:1077-88; Cote, JF et al (1998) Biochemistry 37:13128-37; Nishiya, N (1999) J Biol Chem 274:9847-53; Habib, T et al (1994) J Biol Chem 269:25243-6). Certain PTP-PEST polypeptides are involved in inactivation of the Ras pathway (Davidson, D and Veillette, A (2001) EMBO J 20:3414-26). The expression level of certain PTP-PEST polypeptides can modulate the activity of the GTPase, Rac1 (Sastry, et al (2002) J Cell Sci 115(Pt 22): 4305-16). Certain PTP-PEST polypeptides are involved in the regulation of cell motility (Garton, AJ and Tonks, NK (1999) J Biol Chem 274:3811-8; Angers-Loustau, et al (1999) J Cell Biol 144:1019-31; and Sastry, et al. (2002) J Cell Sci 155(Pt 22): 4305-16). Accordingly, certain POSH polypeptides are involved in inactivation of the Ras pathway. Certain POSH polypeptides are involved in the regulation of cell motility.

Certain PTP-PEST polypeptides are involved in amyloidβ-induced neuronal dystrophy, a pathological hallmark of Alzheimer's disease (Grace, EA and Busciglio, J (2003) J Neurosci. 23:493-502). Accordingly, certain POSH polypeptides may be involved in Alzheimer's disease. Certain PTP-PEST polypeptides function as negative regulators of lymphocyte activation (Davidson, D and Veillette, A (2001) EMBO J 20:3414-26). Accordingly, certain POSH polypeptides may be involved in the regulation of lymphocyte activation. In certain aspects, PTPN12 polypeptides participate in POSH-mediated processes.

The term PTPN12 is used herein to refer to various naturally occurring PTPN12 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring PTPN12 (e.g., SEQ ID NOs: 266-268, 123-129). The term specifically includes human PTPN12 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with HERPUD1, a "homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1" protein. Accordingly, the application provides complexes comprising POSH and HERPUD1. In one aspect, the application relates to the discovery that POSH binds directly with HERPUD1. This interaction was identified by Applicants in a yeast 2-hybrid assay. HERPUD1 is synonymous with Herp. In part, the present application relates to the discovery that a POSH-AP, HERPUD1, is involved in the maturation of an envelope virus, such as HIV.

Certain HERPUD1 polypeptides are involved in JNK-mediated apoptosis, particularly in vascular endothelial cells, including cells that are exposed to high levels of homocysteine. Certain HERPUD1 polypeptides are involved in the Unfolded Protein Response, a cellular response to the presence of unfolded proteins in the endoplasmic reticulum. Certain HERPUD1 polypeptides are involved in the regulation of sterol biosynthesis. Accordingly, certain POSH polypeptides are involved in the Unfolded Protein Response and sterol biosynthesis.

In other aspects, certain HERPUD1 polypeptides enhance presenilin-mediated amyloid β-protein generation. For example, HERPUD1 polypeptides, when overexpressed in cells, increase the level of amyloid β generation, and it is observed that HERPUD1 polypeptides interact with the presenilin proteins, presenilin-1 and presenilin-2. (See Sai, X. et al (2002) J. Biol. Chem. 277:12915-12920). Accordingly, in certain aspects, POSH polypeptides may modulate the level of amyloid β generation. Additionally, POSH polypeptides may interact with presenilin 1 and presenilin 2. Therefore, it is believed certain POSH polypeptides modulate presenilin-mediated amyloid β generation. The accumulation of amyloid

β is one hallmark of Alzheimer's disease. Accordingly, these POSH polypeptides may be involved in the pathogenesis of Alzheimer's disease. At sites such as late intracellular compartment sites including the trans-Golgi network, certain mutant presenilin-2 polypeptides up-regulate production of amyloid β peptides ending at position 42 (Aβ42). (See Iwata, H. et al (2001) J. Biol. Chem. 276: 21678-21685). Accordingly, POSH polypeptides regulate production of Aβ42 through mutant presenilin-2 at late intracellular compartment sites including the trans-Golgi network. Furthermore, elevated homocysteine levels have been found to be a risk factor associated with Alzheimer's disease and cerebral vascular disease. Some risk factors, such as elevated plasma homocysteine levels, may accelerate or increase the severity of several central nervous system (CNS) disorders. Elevated levels of plasma homocysteine were found in young male patients with schizophrenia suggesting that elevated homocysteine levels could be related to the pathophysiology of aspects of schizophrenia (Levine, J. et al (2002) Am. J. Psychiatry 159:1790-2). Accordingly, certain POSH polypeptides may be involved in neurological disorders. Neurological disorders include disorders associated with increased levels of plasma homocysteine, increased levels of amyloid β production, or aberrant presenilin acitivity. Neurological disorders include CNS disorders, such as Alzheimer's disease, cerebral vascular disease and schizophrenia. Certain POSH polypeptides may be involved in cardiovascular diseases, such as thromboembolic vascular disease, and particularly the disease characteristics associated with hyperhomocysteinemia. See, for example, Kokame et al. 2000 J. Biol. Chem. 275:32846-53; Zhang et al. 2001 Biochem Biophys Res Commun 289:718-24.

The term HERPUD1 is used herein to refer to various naturally occurring HERPUD1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring HERPUD1 (e.g., SEQ ID NOs: 249-252, 77-86). The term specifically includes human HERPUD1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with one or more Cbl-b polypeptides. Accordingly, the

application provides complexes comprising POSH and Cbl-b. In one aspect, the application relates to the discovery that POSH binds directly with Cbl-b. This interaction was identified by Applicants in a yeast 2-hybrid assay. Cbl-b polypeptides contain an amino-terminal variant SH2 domain, a RING finger, and a carboxyl-terminal proline-rich domain with potential tyrosine phosphorylation sites. Cbl-b is highly homologous to the mammalian Cbl and the nematode Sli-1 proteins. This application provides four Cbl-b variants and shows that the POSH polypeptide interacts with one or more of these variants. In one aspect, the POSH polypeptide interacts with a human Cbl-b (UniGene No.: Hs.3144). In another aspect, the POSH polypeptide interacts with an alternative human Cbl-b (UniGene No.: Hs.381921) that may be a splice variant of Cbl-b. In yet another aspect, the POSH polypeptide interacts with a human Cbl-b polypeptide that is a splice variant represented by the amino acid sequence depicted in SEQ ID NO: 361, which is encoded by the nucleic acid sequence depicted in SEQ ID NO: 359. In yet another aspect, the POSH polypeptide interacts with a human Cbl-b polypeptide that is a splice variant represented by the amino acid sequence depicted in SEQ ID NO: 398, which is encoded by the nucleic acid sequence depicted in SEQ ID NO: 360.

Certain Cbl-b polypeptides have been shown to function as adaptor proteins by interacting with other signaling molecules, e.g., interaction with cell surface receptor tyrosine kinases, e.g., EGFR (Ettenberg, SA et al (2001) J Biol Chem 276:77-84) or with proteins such as Syk (Elly, C et al (1999) Oncogene 18:1147-56), Crk-L (Elly, C et al (1999) Oncogene 18:1147-56), PI3K (Fang, D et al. (2001) J Biol Chem 16:4872-8), Grb2 (Ettenberg, SA et al (1999) Oncogene 18:1855-66), or Vav (Bustelo, XR et al. (1997) Oncogene 15:2511-20). Certain Cbl-b polypeptides have been demonstrated to interact directly with the nucleotide exchange factor, Vav (Bustelo, XR et al. (1997) Oncogene 15:2511-20). Certain Cbl-b polypeptides have been shown to function as an E3 ubiquitin ligase that recognizes tyrosine phosphorylated substrates through its SH2 domain and through its RING domain, recruits a ubiquitin-conjugating enzyme, E2 (Joazeiro, C et al. (1999) Science 286:309-312) Additionally, certain Cbl-b polypeptides have been shown to associate directly with the p85 subunit of PI3K and to function as an E3 ligase in the ubiquitination of PI3K (Fang, D et al. (2001) J Biol Chem 16:4872-8).

Certain Cbl-b polypeptides are negative regulators of T-cell activation. Cbl-b-deficient mice become very susceptible to experimental autoimmune encephalomyelitis (Chiang, YJ et al. (2000) Nature 403:216-220). Also, Cbl-b-deficient mice develop spontaneous autoimmunity (Bachmaier, K, et al (2000) Nature 403:211-216). Furthermore, Cbl-b is a major susceptibility gene for rat type 1 diabetes mellitus (Yokoi, N et al (2002) Nature Genet. 31:391-394).

Accordingly, in certain aspects, POSH participates in the formation of Cbl-b complexes, including human Cbl-b-containing complexes. Certain POSH polypeptides may be involved in disorders of the immune system, e.g., autoimmune disorders. Certain POSH polypeptides may be involved in the regulation of T-cell activation. In certain aspects, POSH participates in the ubiquitination of PI3K. In certain aspects, Cbl-b polypeptides participate in POSH-mediated processes.

The term Cbl-b is used herein to refer to full-length, human Cbl-b (UniGene No.: Hs.3144) as well as an alternative Cbl-b (UniGene No.: Hs.381921) composed of two separate Cbl-b sequences (e.g., nucleic acid sequences) that may be a splice variant. The term Cbl-b is used herein to refer as well to the human Cbl-b splice variant represented by the amino acid sequence of SEQ ID NO: 361, which is encoded by the nucleic acid sequence of SEQ ID NO: 359 and to the human Cbl-b splice variant represented by the amino acid sequence of SEQ ID NO: 398, which is encoded by the nucleic acid sequence of SEQ ID NO: 360. The term Cbl-b is used herein to refer as well to various naturally occurring Cbl-b homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring Cbl-b (e.g., SEQ ID NOs: 361, 398, 227-230, 353-360). The term specifically includes human Cbl-b nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with GOSR2. Accordingly, the application provides complexes comprising POSH and GOSR2. In one aspect, the application relates to the discovery that POSH binds directly with GOSR2. This interaction was identified by Applicants in a yeast 2-hybrid assay. Certain GOSR2 polypeptides are synonymous with GS27 (for Golgi SNARE of 27K) and are involved in trafficking membrane proteins between the endoplasmic reticulum and the Golgi and between

Golgi subcompartments such as between the cis-,medial- and trans-Golgi network. (See, for example, Lowe, SL et al (1997) Nature 389:881-4 and Bui, TD et al (1999) 57:285-8). Accordingly, certain POSH polypeptides are involved in the trafficking of membrane proteins between the endoplasmic reticulum and the Golgi and between Golgi subcompartments.

The term GOSR2 is used herein to refer to various naturally occurring GOSR2 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring GOSR2 (e.g., SEQ ID NOs: 244-248, 69-76). The term specifically includes human GOSR2 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with RALA. Accordingly, the application provides complexes comprising POSH and RALA. In one aspect, the application relates to the discovery that POSH binds directly with RALA. This interaction was identified by Applicants in a yeast 2-hybrid assay. RALA polypeptides are GTP-binding polypeptides. RALA polypeptides are members of the Ras family of proteins and are GTPases. Certain RALA polypeptides may be synonymous with RalA polypeptides. RalA polypeptides are small GTPases. RalA polypeptides have been shown to interact with phospholipase D and to effect phospholipase D activity. Additionally, RalA polypeptides may be involved in tumor formation and cell transformation. (See, for example, Kim, JH et al (1998) FEBS Lett 430:231-5; Aguirre-Ghiso, JA et al (1999) Oncogene 18:4718-25; Lu, Z et al (2000) Mol Cell Biol 20:462-7; Gildea, JJ et al (2002) Cancer Res 62:982-5; Lucas, L et al (2002) Int J Oncol 21:477-85; and Xu, L et al (2003) Mol Cell Biol 23:645-54). Accordingly, certain POSH polypeptides may interact with PLD and modulate its activity, and certain POSH polypeptides may be involved in tumor formation and cell transformation. In other aspects, certain RalA polypeptides interact with calmodulin and may be involved in calcium/calmodulin-mediated intracellular signaling pathways (Clough, RR et al (2002) J Biol Chem 277:28972-80). Certain RalA polypeptides are involved in controlling actin cytoskeletal remodeling and vesicle transport in mammalian cells. Certain RalA polypeptides interact with the exocyst complex, which is involved in exocytosis. (See, for example, Sugihara, K et al (2002) Nat Cell Biol 4:73-8; Polzin,

A et al (2002) Mol Cell Biol 22:1714-22; and Lipschutz, JH and Mostov, KE (2002) Curr Biol 12(6):R212-4). Accordingly, certain POSH polypeptides are involved in vesicle transport.

The term RALA is used herein to refer to various naturally occurring RALA homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring RALA (e.g., SEQ ID NOs: 269-270, 130-134). The term specifically includes human RALA nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with SMN1. Accordingly, the application provides complexes comprising POSH and SMN1. In one aspect, the application relates to the discovery that POSH binds directly with SMN1. This interaction was identified by Applicants in a yeast 2-hybrid assay. SMN1 polypeptides are encoded by the nucleic acid of the survival motor neuron gene 1 (SMN1). Mutations in this gene (such as its homozygous absence) cause spinal muscular atrophy (SMA), a common autosomal recessive disorder characterized by degeneration of motor neurons in the spinal cord, leading to progressive paralysis with muscular atrophy. Accordingly, POSH may be involved in the pathogenesis of SMA. SMN1 is part of a multiprotein complex that is required for biogenesis of the Sm class of small nuclear ribonucleoproteins (Sm snRNPs). SMN1 associates with a number of proteins, such as Gemin2 to Gemin6, to form a large complex found in both the cytoplasm and in the nucleus. SMN1 also associates with Snurportin 1, an adaptor protein that recognizes the nuclear localization signal of Sm snRNPs. (See, for example, Lefebvre, S et al (1995) Cell 80:155-65; Narayanan, U et al (2002) Hum Mol Genet 11:1785-95; Massenet, S et al (2002) 22:6533-41; and Monani, UR et al (1999) Hum Mol Genet 8:1177-83). Accordingly, certain POSH polypeptides may be involved in the biogenesis of snRNPs. Certain SMN1 polypeptides interact with the large nonstructural protein NS1 of the autonomous parvovirus minute virus of mice (MVM). NS1 is essential for viral replication, and it is a potent transcriptional activator (Young, PJ et al (2002) J Virol 76:3892-904). Certain SMN1 polypeptides interact with the protein NS2 of MVM. NS2 is also required for efficient viral replication. Certain SMN1 polypeptides colocalize with NS2 in infected nuclei and at late times following

MVM infection. (See Young, PJ et al (2002) J Virol 76:6364-9). Accordingly, POSH polypeptides are involved in viral replication.

_The term SMN1 is used herein to refer to various naturally occurring SMN1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring SMN1 (e.g., SEQ ID NOs: 273-275, 142-146). The term specifically includes human SMN1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with SMN2. Accordingly, the application provides complexes comprising POSH and SMN2. In one aspect, the application relates to the discovery that POSH binds directly with SMN2. This interaction was identified by Applicants in a yeast 2-hybrid assay. The SMN2 gene is an almost identical copy of the SMN1 gene that causes SMA. A critical difference between the two genes is a 1 nucleotide base change inside exon 7 that affects the splicing pattern of the genes. The majority of the SMN2 transcript lacks exon 7. Certain SMN2 polypeptides influence the severity of SMA. (See, for example, Monani, UR et al (1999) Hum Mol Genet 8: 1177-83; Cartegni, L and Krainer, AR (2002) Nat Genet 30:377-84; and Feldkotter, M et al (2002) Am J Hum Genet 70: 358-68). Accordingly, certain POSH polypeptides may influence the severity of SMA.

The term SMN2 is used herein to refer to various naturally occurring SMN2 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring SMN2 (e.g., SEQ ID NOs: 276-280, 147-151). The term specifically includes human SMN2 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with SIAH1. Accordingly, the application provides complexes comprising POSH and SIAH1. In one aspect, the application relates to the discovery that POSH binds directly with SIAH1. This interaction was identified by Applicants in a yeast 2-hybrid assay. Certain SIAH1 polypeptides bind ubiquitin-conjugating enzymes and target proteins for proteasome-mediated degradation. Certain SIAH1 polypeptides are involved in targeting beta-catenin for degradation (Matsuzawa, S and Reed, JC (2001) Molec Cell 7: 915-926 and Liu, J et al (2001) Molec Cell 7:

927-936). Accordingly, certain POSH polypeptides are involved in the targeting of beta-catenin for degradation. Certain SIAH1 polypeptides are E3 ubiquitin ligases and regulate the ubiquitination and degradation of synaptophysin (Wheeler, TC et al. (2002) J Biol Chem 277: 10273-92). Accordingly, certain POSH polypeptides are involved in the ubiquitination and degardation of synaptophysin. Certain SIAH1 polypeptides regulate the protein, DCC (deleted in colorectal cancer), via the ubiquitin-proteosome pathway (Hu, G et al. (1997) Genes Dev 11: 2701-14). Accordingly, certain POSH polypeptides are involved in the ubiquitination and degardation of DCC. Certain SIAH1 polypeptides are a target of activation of p53 and are upregulated by p53, and certain SIAH1 polypeptides are involved in apoptosis, tumor suppression, as well as vertebrate development (Maeda, A et al (2002) FEBS Lett 512: 223-226; Hu, G et al (1997) Genomics 46:103-111; and Nemani, M et al (1996) Proc Natl Acad Sci USA 93: 9039-9042). Accordingly, certain POSH polypeptides may be a target of p53 activation, and certain POSH polypeptides may be involved in apoptosis and tumor suppression.

The term SIAH1 is used herein to refer to various naturally occurring SIAH1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring SIAH1 (e.g., SEQ ID NOs: 271-272, 135-141). The term specifically includes human SIAH1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with SYNE1. Accordingly, the application provides complexes comprising POSH and SYNE1. In one aspect, the application relates to the discovery that POSH binds directly with SYNE1. This interaction was identified by Applicants in a yeast 2-hybrid assay. SYNE1 polypeptides are synonymous with Syne-1, myne-1, and nesprin-1 polypeptides. Syne-1 polypeptides are associated with nuclear envelopes in skeletal, cardiac, and smooth muscle cells. Syne-1 polypeptides contain multiple spectrin repeats. In muscle, myne-1 expression is observed in the inner nuclear envelope, and myne-1 has been shown to interact with the inner nuclear membrane protein lamin A/C. Syne-1 also associates with the nuclear envelope protein, emerin. Syne-1 polypeptides may be involved in maintaining nuclear organization and structural integrity, and certain Syne-1

polypeptides may be involved in the migration of myonuclei in myotubes and/or their anchoring at the postsynaptic apparatus. (See, for example, Apel et al (2000) J Biol Chem 275:31986-95; Zhang, Q et al (2001) J Cell Sci 114:4485-98; Zhang, Q et al (2002) Genomics 80:473-81; and Mislow, JM et al (2002) J Cell Sci 115 (Pt 1):61-70). Accordingly, certain POSH polypeptides may interact with the lamin A/C polypeptides and/or emerin polypeptides. Also, certain POSH polypeptides may be involved in maintaining nuclear organization and structural integrity, and certain POSH polypeptides may be involved in the migration of myonuclei in myotubes and/or their anchoring at the postsynaptic apparatus.

The term SYNE1 is used herein to refer to various naturally occurring SYNE1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring SYNE1 (e.g., SEQ ID NOs: 295-307, 183-201). The term specifically includes human SYNE1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with TTC3. Accordingly, the application provides complexes comprising POSH and TTC3. In one aspect, the application relates to the discovery that POSH binds directly with TTC3 This interaction was identified by Applicants in a yeast 2-hybrid assay. Certain TTC3 polypeptides are synonymous with the proteins, TPRDI, TPRDII, TRPDIII, TPRD and DCRR1 and may be involved in the pathogenesis of certain characteristics of Down syndrome, such as morphological features, hypotonia, and mental retardation (Tsukahar, F et al (1996) J Biochem (Tokyo) 120: 820-827; Ohira, M et al (1996) DNA Res 3: 9-16; Dahmane, N et al (1998) Genomics 48: 12-23; and Eki, T et al (1997) DNA Seq 7:153-164).

The term TTC3 is used herein to refer to various naturally occurring TTC3 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring TTC3 (e.g., SEQ ID NOs: 308-312, 202-207). The term specifically includes human TTC3 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with VCY2IP1. Accordingly, the application provides

complexes comprising POSH and VCY2IP1. In one aspect, the application relates to the discovery that POSH binds directly with VCY2IP1. This interaction was identified by Applicants in a yeast 2-hybrid assay. VCY2IP1 is synonymous with VCY2IP-1, which has been shown to interact with the testis-specific protein, VCY2. VCY2IP1 is also synonymous with C19orf5, which has been shown to interact with the tumor suppressor, RASSF1, suggesting a role for C19orf5 in apoptosis and tumor suppression (In Vitro Cell Dev Biol Anim (2002) 38:582-94). C19orf5 also demonstrates a strong homology to microtubule-associated proteins (Genomics (2002) 79:124-6). Accordingly, POSH may play a role in apoptosis and tumor suppression.

The term VCY2IP1 is used herein to refer to various naturally occurring VCY2IP1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring VCY2IP1 (e.g., SEQ ID NOs: 315-323, 214-222). The term specifically includes human VCY2IP1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with MSTP028. In one aspect, the application relates to the discovery that POSH binds directly with MSTP028. This interaction was identified by Applicants in a yeast 2-hybrid assay. In part, the present application relates to the discovery that a POSH-AP, MSTP028, is involved in the maturation of an envelope virus, such as HIV. Certain MSTP028 polypeptides contain one or more BTB/POZ domains that are generally involved in dimerization. Accordingly the application provides complexes comprising POSH and MSTP028, optionally in a dimeric form. The term MSTP028 is used herein to refer to various naturally occurring MSTP028 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring MSTP028 (e.g., SEQ ID NOs: 255-256, 90-94). The term specifically includes human MSTP028 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with SNX1. Accordingly, the application provides complexes comprising POSH and SNX1. In one aspect, the application relates to the discovery

that POSH binds directly with SNX1. This interaction was identified by Applicants in a yeast 2-hybrid assay. SNX1 is a member of the sorting nexin (SNX) protein family, which is implicated in regulating membrane traffic. SNX1 is a membrane associated protein that has been shown to be involved with targeting receptors to lysosomal degradation. SNX1 has been shown to bind to the C-terminal tail of the D5 dopamine receptor (Mol Cell Biol (1998) 18: 7278-87). Accordingly, in certain aspects POSH may associate with the D5 dopamine receptor. SNX1 is involved in regulating the targeting of internalized epidermal growth factor receptors for lysosomal degradation (Science (1996) 272:1008-1010). In certain aspects, POSH may be involved in targeting proteins for degradation to the lysosome. SNX1 has also been found to be involved in sorting PAR1, a G-protein coupled receptor for thrombin (Mol Cell Biol (2002) 13:1965-76). It has further been demonstrated that SNX1 functions in regulating trafficking in the endosome compartment via recognition of phosphorylated phosphatidylinositol through the phox homology domain (PX domain) of SNX1 (Proc Natl Acad Sci (2002) 99:6767-72).

The term SNX1 is used herein to refer to various naturally occurring SNX1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring SNX1 (e.g., SEQ ID NOs: 281-286, 152-161). The term specifically includes human SNX1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In additional embodiments, the application relates to the discovery that a POSH polypeptide interacts with SNX3. Accordingly, the application provides complexes comprising POSH and SNX3. In one aspect, the application relates to the discovery that POSH binds directly with SNX3. This interaction was identified by Applicants in a yeast 2-hybrid assay. SNX3 is also a member of the SNX protein family. SNX3 has been shown to associate with the early endosome through its PX domain, a domain capable of interaction with phosphatidylinositol-3-phosphate (Nat Cell Biol (2002) 3:658-66). Accordingly, POSH may be involved in membrane traffic at the early endosome.

The term SNX3 is used herein to refer to various naturally occurring SNX3 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring SNX3 (e.g., SEQ

ID NOS: 287-290, 162-174). The term specifically includes human SNX3 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In further embodiments, the application relates to the discovery that a POSH polypeptide interacts with ATP6V0C. Accordingly, the application provides complexes comprising POSH and ASTP6V0C. In one aspect, the application relates to the discovery that POSH binds directly with ATP6V0C. This interaction was identified by Applicants in a yeast 2-hybrid assay. ATP6V0C, vacuolar-H(+)-ATPase, is a large multimeric protein composed of at least twelve distinct subunits and it is involved in the H(+) transport across cellular membranes. ATP6V0C is synonymous with ATP6L. Treatment with anticancer agents has been shown to enhance ATP6L expression (Cytogenet Genome Res (2002) 97:111-5; J Biol Chem (2002) 277:36534-43).

The term ATP6V0C is used herein to refer to various naturally occurring ATP6V0C homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring ATP6V0C (e.g., SEQ ID NOs: 225-226, 345-351). The term specifically includes human ATP6V0C nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with PPP1CA. Accordingly, the application provides complexes comprising POSH and PPP1CA. In one aspect, the application relates to the discovery that POSH binds directly with PPP1CA. This interaction was identified by Applicants in a yeast 2-hybrid assay. PPP1CA is the protein phosphatase type 1 alpha catalytic subunit. The genetic and expression status of the PPP1CA gene was examined in 55 human cancer cell lines and found to be ubiquitously expressed and lacking in genetic variation, suggesting an essential role for PPP1CA in the growth of cancer cells (Int J Oncol (2001) 18:817-24).

The term PPP1CA is used herein to refer to various naturally occurring PPP1CA homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring PPP1CA (e.g., SEQ ID NOs: 261-263, 101-110). The term specifically includes human

PPP1CA nucleic acid and amino acid sequences and the sequences presented in Figure 36.

The application further relates to the discovery that a POSH polypeptide interacts with DDEF1. Accordingly, the application provides complexes comprising POSH and DDEF1. In one aspect, the application relates to the discovery that POSH binds directly with DDEF1. This interaction was identified by Applicants in a yeast 2-hybrid assay. DDEF1 is a putative candidate gene associated with Meckel-Gruber syndrome (MKS), the most common monogenic cause of neural tube defects (Hum Genet (2002) 111:654-61).

The term DDEF1 is used herein to refer to various naturally occurring DDEF1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring DDEF1 (e.g., SEQ ID NOs: 233-237, 48-54). The term specifically includes human DDEF1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with PACS-1. Accordingly, the application provides complexes comprising POSH and PACS-1. In one aspect, the application relates to the discovery that POSH binds directly with PACS-1. This interaction was identified by Applicants in a yeast 2-hybrid assay. PACS-1 is a cytosolic sorting protein that directs localization of membrane proteins in the TGN/endosomal system. PACS-1 is a cytosolic protein involved in controlling the correct subcellular localization of integral membrane proteins that contain acidic cluster sorting motifs, such as furin and HIV-1 Nef, and PACS-1 has been shown to interact with the adaptor complexes AP-1 and AP-3 (EMBO J (2003) 22:6234-44; EMBO J (2001) 20:2191-201). Furthermore, PACS-1 polypeptides have been shown to interact with Nef and through this interaction, by a PI3K-dependent proces, MHC class I molecules are downregulated by Nef (Cell (2002) 11:853-66). Accordingly, POSH may be involved in Nef-mediated downregulation of MHC class I molecules in a cell infected with HIV-1. Additionally, PACS-1 interacts with the HIV-1 protein, Vpu. Vpu expresses an acidic amino acid sorting motif that is required for TGN localization through a retroviral process mediated by PACS-1 (Wan, L et al (1998)

Cell 94:205-216). Accordingly, in certain aspects, POSH may associate with Vpu through its interaction with PACS-1.

The term PACS-1 is used herein to refer to various naturally occurring PACS-1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring PACS-1 (e.g., SEQ ID NOs: 362-366, 95-100). The term specifically includes human PACS-1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with EPS8L2. Accordingly, the application provides complexes comprising POSH and EPS8L2. In one aspect, the application relates to the discovery that POSH binds directly with EPS8L2. This interaction was identified by Applicants in a yeast 2-hybrid assay. EPS8L2 is an eps8-related protein. Eps8 forms a multimeric complex with Sos-1, Abi1 and PI3K that is required for Rac activation leading to actin remodelling. EPS8L2 has been shown to interact with Abi1 and Sos-1. EPS8L2 also has been shown to localize to PDGF-induced F-actin-rich ruffles and to restore receptor tyrosine kinase mediated actin remodeling when expressed in eps8-/- fibroblasts (Mol Biol Cell (2004) 15:91-8).

The term EPS8L2 is used herein to refer to various naturally occurring EPS8L2 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring EPS8L2 (e.g., SEQ ID NOs: 239, 58-60). The term specifically includes human EPS8L2 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

The application additionally relates to the discovery that a POSH polypeptide interacts with HIP55. Accordingly, the application provides complexes comprising POSH and HIP55. In one aspect, the application relates to the discovery that POSH binds directly with HIP55. This interaction was identified by Applicants in a yeast 2-hybrid assay. HIP55 is a cytoplasmic adaptor protein that has been shown to bind to the cytoplasmic tail of the CD2v protein of African swine fever virus (J Gen Virol (2004) 85:119-30). HIP55 (synonymous with mAbp1 and SH3P7) comprises an SH3 domain and through its SH3 domain, associates with dynamin (J Cell Biol (2001) 153:351-66; Biochem Biophys Res Commun (2003) 301:704-10). Accordingly, in certain aspects, POSH may associate with dynamin through its

interaction with HIP55. HIP55 has also been shown to be important for receptor mediated endocytosis of the transferrin receptor (Biochem Biophys Res Commun (2003) 301:704-10).

The term HIP55 is used herein to refer to various naturally occurring HIP55 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring HIP55 (e.g., SEQ ID NOs: 390-394, 377-385). The term specifically includes human HIP55 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with CENTB1. Accordingly, the application provides complexes comprising POSH and CENTB1. In one aspect, the application relates to the discovery that POSH binds directly with CENTB1. This interaction was identified by Applicants in a yeast 2-hybrid assay. CENTB1 is synonymous with ACAP1. ACAP1 is an ARF GTPase activating protein (ARF GAP). ACAP1 can function as a GAP for ARF1 and ARF6 (J Biol Chem (2002) 277:7962-9).

The term CENTB1 is used herein to refer to various naturally occurring CENTB1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring CENTB1 (e.g., SEQ ID NOs: 231-232, 37-47). The term specifically includes human CENTB1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with EIF3S3. Accordingly, the application provides complexes comprising POSH and EIF3S3. In one aspect, the application relates to the discovery that POSH binds directly with EIF3S3. This interaction was identified by Applicants in a yeast 2-hybrid assay. EIF3S3 is elevated in certain hepatocellular carcinomas and in prostate cancer (Hepatology (2003) 38:1242-9; Am J Pathol (2001) 159:2081-84). It has also been demonstrated that EIF3S3 is often amplified and overexpressed in breast cancer (Genes Chromosomes Cancer. (2000) 28:203-210).

The term EIF3S3 is used herein to refer to various naturally occurring EIF3S3 homologs, as well as functionally similar variants and fragments that retain

at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring EIF3S3 (e.g., SEQ ID NOs: 238, 55-57). The term specifically includes human EIF3S3 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the discovery that a POSH polypeptide interacts with SRA1. Accordingly, the application provides complexes comprising POSH and SRA1. In one aspect, the application relates to the discovery that POSH binds directly with SRA1. This interaction was identified by Applicants in a yeast 2-hybrid assay. SRA1 is a transcriptional coactivator, steroid receptor RNA activator 1. SRA is selective for steroid hormone receptors and mediates transactivation via their amino-terminal activation function (Cell (1999) 97:17-27). The term SRA1 is used herein to refer to various naturally occurring SRA1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring SRA1 (e.g., SEQ ID NOs: 291-294, 175-182). The term specifically includes human SRA1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

The application additionally relates to the discovery that a POSH polypeptide interacts with WASF1. Accordingly, the application provides complexes comprising POSH and WASF1. In one aspect, the application relates to the discovery that POSH binds directly with WASF1. This interaction was identified by Applicants in a yeast 2-hybrid assay. WASF1 is a member of the Wiskott-Aldrich syndrome protein (WASP) family of proteins. WASF-1 has been shown to regulate cortical actin filament reorganization in response to extracellular stimuli. WASF1 is synonymous with WAVE1 and is an actin regulatory protein. It has been shown that Ras and the adaptor protein Nck activate actin nucleation through WAVE1 (Nature (2002) 418:790-3).

The term WASF1 is used herein to refer to various naturally occurring WASF1 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring WASF1 (e.g., SEQ ID NOs: 389, 375-376). The term specifically includes human WASF1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

The application additionally relates to the discovery that a POSH polypeptide interacts with SPG20. Accordingly, the application provides complexes comprising

POSH and SPG20. In one aspect, the application relates to the discovery that POSH binds directly with SPG20. This interaction was identified by Applicants in a yeast 2-hybrid assay. SPG20 is synonymous with spartin, and mutation in the gene has been implicated in Troyer syndrome, an autosomal recessive complicated hereditary spastic paraplegia. Comparative sequence analysis has shown that spartin shares similarity with molecules involved in endosomal trafficking (Nat Genet (2002) 31:347-8).

The term SPG20 is used herein to refer to various naturally occurring SPG20 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring SPG20 (e.g., SEQ ID NOs: 386-388, 367-374). The term specifically includes human SPG20 nucleic acid and amino acid sequences and the sequences presented in the Figure 36.

In further embodiments, the application relates to the discovery that a POSH polypeptide interacts with HLA-A. Accordingly, the application provides complexes comprising POSH and HLA-A. In one aspect, the application relates to the discovery that POSH binds directly with HLA-A. This interaction was identified by Applicants in a yeast 2-hybrid assay. In additional aspects, the application relates to the discovery that a POSH polypeptide interacts with HLA-B. Accordingly, the application provides complexes comprising POSH and HLA-B. In one aspect, the application relates to the discovery that POSH binds directly with HLA-B. This interaction was identified by Applicants in a yeast 2-hybrid assay. HLA-A and HLA-B are MHC class I molecules. HLA-A and HLA-B molecules are downregulated in the progression of AIDS, and this downregulation is associated with the activity of HIV-1 Nef.

The term HLA-A is used herein to refer to various naturally occurring HLA-A homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring HLA-A (e.g., SEQ ID NOs: 253, 87-88). The term specifically includes human HLA-A nucleic acid and amino acid sequences and the sequences presented in Figure 36.

The term HLA-B is used herein to refer to various naturally occurring HLA-B homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring HLA-B

(e.g., SEQ ID NOs: 254, 89). The term specifically includes human HLA-B nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain aspects, the application relates to the discovery that a POSH polypeptide interacts with a ubiquitin-conjugating enzyme (E2). An exemplary E2 may include, but are not limited to, UBC5a, UBC5c, UBC6, and UBC13. UBC13 is often found in a heterodimer complex with a Ub conjugating enzymer variant (UEV) protein, such as, for example, UEV1a. (See Hofmann and Pickart, *Noncanonical MMS2-Encoded Ubiquitin-Conjugating Enzyme Functions in Assembly of Novel Ubiquitin Chains for DNA Repair, Cell* 96: 645-653 (1999), McKenna et al., 2002, *Energetics and Specificity of Interactions within Ub-Uev-Ubc13 Human Ubiquitin Conjugating Complexs, Biochemistry.* Vol. 42. pp.7922-7930, and Ulrich, 2003, *Protein-Protein Interactions within an E2-RING Finger Complex, The Jurnal of Biological Chemistry,* Vol. 278. No 9. pp. 7051-7058). UVE proteins share significant sequence and structural similarities with E2s, yet lack the requisite active site cystine of the classical E2 protein family.

Generally, UBC5 conjugates ubiquitin to Lysine 48 in a target protein, a signal that marks the protein for degredation by the 26 S proteosome. In constrast, UBC13/UEV1a conjugates ubiquitin to Lysine 63 residue in a target protein, which is not a degradation signal. Instead, ubiquitin conjugated at Lysine 63 has been implicated in diverse biological processes, including, for example, DNA damage repair, endocytosis, ribosome biogenesis, mitochondrial inheritance, and NFκB signaling (See Ulrich, 2003). The UBC13/UEV1a has been shown to work with two other RING-ubiquitin ligases, TRAF6 and RAD5. (See Ulrich, 2003). TRAF6-UBC13-UEV1a complex ubiquitinates TRAF6 (self-ubiquitination), thus enabling it to activate a kinase cascade.

Without being bound to theory, it appears that UBC5a, UBC5c and UBC6 may work with POSH in one pathway, while UBC13/UEV1a work with POSH in another distinct pathway. This is supported by the fact that UBC5/6 marks POSH for degradation by conjugating ubiquitin at Lysine 48, whereas UBC13/UEV1a marks POSH for purposes other than degradation by conjugating ubiquitin at Lysine 63. This theory is further supported by the fact that UBC5a, UBC5c and UBC6 share high sequence similarities.

Accordingly, in certain aspects, the present application relates to an isolated, purified or recombinant complex comprising a POSH polypeptide and a UBC13. In certain aspects, the present application relates to an isolated, purified or recombinant complex comprising: a polypeptide comprising a domain that is at least 90% identical to a POSH RING domain, and a POSH-AP comprising an E2. An exemplary POSH associated protein E2 include, for example, is UBC13. UBC13 may be in a heterodimer complex with a Ub conjugating enzymer variant (UEV) protein, such as, for example, UEV1a.

The term "UBC13" and is used herein to refer to full-length UBC13, any splice variants thereof, various naturally occurring UBC13 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring UBC13 (e.g., SEQ ID NOs: 313, 208-210). The term specifically includes UBC13 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the interaction between an ARF5 polypeptide and a POSH polypeptide. ARF5 is a member of the ARF gene family. The ARF proteins stimulate the in vitro ADP-ribosyltransferase activity of cholera toxin. ARF proteins play a role in vesicular trafficking in vivo. ARFs are members of the Ras GTPase superfamily. ARFs activate specific PLDs. Mammalian ARFs are divided into three classes based on size, amino acid sequence, gene structure, and phylogenetic analysis. ARF1 is in class I, and ARF5 is in class II.

In certain embodiments, the application relates to the interaction between an ARF1 polypeptide and a POSH polypeptide. ARF1 is a small G protein involved in vesicular trafficking. The assembly/disassembly cycle of the coat protein I (COPI) on Golgi membranes is coupled to the GTP/GDP cycle of ARF1 (Nature (2003) 426:563-6). ARF1 has been implicated in mitotic Golgi disassembly, chromosome segregation, and cytokinesis (Proc Natl Acad Sci (2003) 100:13314-9). ARF1 has been shown to bind to the 5-HT2A receptor, a G protein coupled receptor (GPCR) (Mol Pharmacol (2003) 64:1239-50).

The term ARF-1 is used herein to refer to various naturally occurring ARF-1 homologs, as well as functionally similar variants and fragments that retain at least

80%, 90%, 95%, or 99% sequence identity to a naturally occurring ARF-1 (e.g. SEQ ID NOs: 223, 325-339). The term specifically includes human ARF-1 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

The term ARF-5 is used herein to refer to various naturally occurring ARF-5 homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring ARF-5 (e.g., SEQ ID NOs: 224, 340-344). The term specifically includes human ARF-5 nucleic acid and amino acid sequences and the sequences presented in Figure 36.

In certain embodiments, the application relates to the inhibition of viral maturation by modulation of an activity associated with a dynamin II polypeptide. Dynamin II is a large GTP-binding protein that is involved in endocytosis and in vesicle formation at the trans-Golgi network. Dynamin II contains a pleckstrin homology domain (PHD) and a proline-rich domain (PRD). Dynamin II plays an important role in vesicle formation at the plasma membrane, trans-Golgi network, and various other intracellular organelles. Accordingly, disrupting the activity of a dynamin II polypeptide or the interaction between a POSH polypeptide and a dynamin II polypeptide (e.g., by reducing POSH protein levels or alternatively, reducing dynamin II protein levels, through RNAi) may disrupt the activity of dynamin II in the secretory pathway and prevent the secretion of viral proteins, such as, for example, HBV proteins. Dynamin II participates in the transport and secretion of HBV proteins (Abdulkarim, AS et al (2003) J. Hepat. 38:76-83). Accordingly, in certain embodiments, inhibition of POSH adversely effects the transport and release of HBV proteins.

In certain embodiments, the application relates to the inhibition of dynamin activity, in particular the inhibition of the activity of dynamin II, a member of the dynamin family of proteins. In certain embodiments, the application relates to inhibition of dynamin II activity, which inhibition disrupts the transport and secretion of HBV proteins. The term dynamin II is used herein to refer to full-length, human dynamin II as well as various naturally occurring dynamin II homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring dynamin II (e.g.,

public gi number: 1196422, public gi number: 1706539, public gi number: 1196423, and public gi number: 1363934).

In certain embodiments, the application relates to the inhibition of viral maturation by modulation of an activity associated with a Vpu polypeptide. Vpu is an HIV-1 encoded ion channel, which, among other tasks in the HIV-1 life cycle, is necessary for efficient virus budding (Schubert, U et al (1995) J. Virol. 69:7699-7711). Vpu may function at the trans Golgi network (TGN). Vpu expresses an acidic amino acid sorting motif that is required for TGN localization through a retroviral process mediated by the POSH-AP, PACS-1 (Wan, L et al (1998) Cell 94:205-216). Moreover, the phenotype conferred by human POSH knockdown is similar to that observed in cells expressing HIV-1 lacking Vpu where viruses also accumulate in intracellular membranes (Klimkait, T et al (1990) J. Virol. 64:621-629).

Vpu regulates virus release from a post-endoplasmic reticulum compartment, such as possibly the TGN, by an ion channel activity mediated by its transmembrane anchor. Vpu also induces the selective down regulation of host cell receptor proteins such as CD4 and major histocompatibility complex class I molecules, in a process involving its cytoplasmic tail. Furthermore, Vpu-mediated degradation of CD4 is dependent on an intact ubiquitin-conjugating system. (See Schubert, U et al (1998) J. Virol. 72:2280-8). In certain embodiments of the present invention, Vpu-mediated degradation of a protein such as CD4 may involve a ubiquitin-conjugating system that includes a POSH polypeptide or a POSH-AP, such as, for example, Cbl-b.

Vpu nucleic acid and the corresponding amino acid sequence encoded thereby are exemplified by the Vpu discussed in Strebel, K et al (1988) 241:1221-1223. The term Vpu is used herein to refer as well to Vpu of other HIV-1 isolates, such as the Vpu disclosed in GenBank, accession number U51190, and the Vpu disclosed in GenBank, accession number U52953. The term Vpu is used herein to refer as well to various naturally occurring Vpu homologs, as well as functionally similar variants and fragments that retain at least 80%, 90%, 95%, or 99% sequence identity to a naturally occurring Vpu.

4. Methods and Compositions for Treating POSH-associated Diseases

In certain aspects, the application provides methods and compositions for treatment of POSH-associated diseases (disorders), including cancer and viral disorders, as well as disorders associated with unwanted apoptosis, including, for example a variety of neurodegenerative disorders, such as Alzheimer's disease.

In certain embodiments, the application relates to viral disorders (e.g., viral infections), and particularly disorders caused by retroid viruses, RNA viruses and/or envelope viruses. In view of the teachings herein, one of skill in the art will understand that the methods and compositions of the application are applicable to a wide range of viruses such as, for example, retroid viruses, RNA viruses, and envelope viruses. In a preferred embodiment, the present application is applicable to retroid viruses. In a more preferred embodiment, the present application is further applicable to retroviruses (retroviridae). In another more preferred embodiment, the present application is applicable to lentivirus, including primate lentivirus group. In a most preferred embodiment, the present application is applicable to Human Immunodeficiency virus (HIV), Human Immunodeficiency virus type-1 (HIV-1), Hepatitis B Virus (HBV) and Human T-cell Leukemia Virus (HTLV).

While not intended to be limiting, relevant retroviruses include: C-type retrovirus which causes lymphosarcoma in Northern Pike, the C-type retrovirus which infects mink, the caprine lentivirus which infects sheep, the Equine Infectious Anemia Virus (EIAV), the C-type retrovirus which infects pigs, the Avian Leukosis Sarcoma Virus (ALSV), the Feline Leukemia Virus (FeLV), the Feline Aids Virus, the Bovine Leukemia Virus (BLV), Moloney Murine Leukemia Virus (MMuLV), the Simian Leukemia Virus (SLV), the Simian Immuno-deficiency Virus (SIV), the Human T-cell Leukemia Virus type-I (HTLV-I), the Human T-cell Leukemia Virus type-II (HTLV-II), Human Immunodeficiency virus type-2 (HIV-2) and Human Immunodeficiency virus type-1 (HIV-1).

The method and compositions of the present application are further applicable to RNA viruses, including ssRNA negative-strand viruses and ssRNA positive-strand viruses. The ssRNA positive-strand viruses include Hepatitis C Virus (HCV). In a preferred embodiment, the present application is applicable to mononegavirales, including filoviruses. Filoviruses further include Ebola viruses

and Marburg viruses. In another preferred embodiment, the present invention is applicable to flaviviruses, including West Nile Virus (WNV).

Other RNA viruses include picornaviruses such as enterovirus, poliovirus, coxsackievirus and hepatitis A virus, the caliciviruses, including Norwalk-like viruses, the rhabdoviruses, including rabies virus, the togaviruses including alphaviruses, Semliki Forest virus, denguevirus, yellow fever virus and rubella virus, the orthomyxoviruses, including Type A, B, and C influenza viruses, the bunyaviruses, including the Rift Valley fever virus and the hantavirus, the filoviruses such as Ebola virus and Marburg virus, and the paramyxoviruses, including mumps virus and measles virus. Additional viruses that may be treated include herpes viruses.

The methods and compositions of the present application are further applicable to hepatotrophic viruses, including HAV, HBV, HCV, HDV, and HEV. In certain aspects, the application relates to a method of inhibiting a hepatotrophic virus, comprising administering a POSH inhibitor to a subject in need thereof. In further aspects, the application relates to a method of treating a viral hepatitis infection, comprising administering a POSH inhibitor to a subject in need thereof. A viral hepatitis infection may be caused by a hepatotrophic virus, such as HAV, HBV, HCV, HDV, or HEV. In certain embodiments, the application relates to a method of treating an HBV infection by administering a POSH inhibitor to a subject in need thereof.

In other embodiments, the application relates to methods of treating or preventing cancer diseases. The terms "cancer," "tumor," and "neoplasia" are used interchangeably herein. As used herein, a cancer (tumor or neoplasia) is characterized by one or more of the following properties: cell growth is not regulated by the normal biochemical and physical influences in the environment; anaplasia (e.g., lack of normal coordinated cell differentiation); and in some instances, metastasis. Cancer diseases include, for example, anal carcinoma, bladder carcinoma, breast carcinoma, cervix carcinoma, chronic lymphocytic leukemia, chronic myelogenous leukemia, endometrial carcinoma, hairy cell leukemia, head and neck carcinoma, lung (small cell) carcinoma, multiple myeloma, non-Hodgkin's lymphoma, follicular lymphoma, ovarian carcinoma, brain tumors, colorectal

carcinoma, hepatocellular carcinoma, Kaposi's sarcoma, lung (non-small cell carcinoma), melanoma, pancreatic carcinoma, prostate carcinoma, renal cell carcinoma, and soft tissue sarcoma. Additional cancer disorders can be found in, for example, Isselbacher et al. (1994) Harrison's Principles of Internal Medicine 1814-1877, herein incorporated by reference.

In a specific embodiment, anticancer therapeutics of the application are used in treating a POSH-associated cancer. As described herein, POSH-associated cancers include, but are not limited to, the thyroid carcinoma, liver cancer (hepatocellular cancer), lung cancer, cervical cancer, ovarian cancer, renal cell carcinoma, lymphoma, osteosacoma, liposarcoma, leukemia, breast carcinoma, and breast adeno-carcinoma.

Preferred antiviral and anticancer therapeutics of the application can function by disrupting the biological activity of a POSH polypeptide or POSH complex in viral maturation. Certain therapeutics of the application function by disrupting the activity of a POSH-AP (e.g., HERPUD1) in viral maturation. Certain therapeutics of the application function by disrupting the activity of POSH by inhibiting the ubiquitin ligase activity of a POSH polypeptide. In certain embodiments of the application, a therapeutic of the application inhibits the ubiquitination of a POSH-AP, such as for example the ubiquitination of HERPUD1.

In other embodiments, the application relates to methods of treating or preventing neurological disorders. In one aspect, the invention provides methods and compositions for the identification of compositions that interfere with the function of a POSH or a POSH-AP, which function may relate to aberrant protein processing associated with a neurodegenerative disorder, such as for example, the processing of amyloid beta precursor protein associated with Alzheimer's disease. Neurological disorders include disorders associated with increased levels of amyloid β production, such as for example, Alzheimer's disease. Neurological disorders also include Parkinson's disease, Huntington's disease, schizophrenia, Niemann-Pick's disease, and prion-associated diseases

Exemplary therapeutics of the application include nucleic acid therapies such as, for example, RNAi constructs (small inhibitory RNAs), antisense

oligonucleotides, ribozyme, and DNA enzymes. Other therapeutics include polypeptides, peptidomimetics, antibodies and small molecules.

Antisense therapies of the application include methods of introducing antisense nucleic acids to disrupt the expression of POSH polypeptides or proteins that are necessary for POSH function.

RNAi therapies include methods of introducing RNAi constructs to downregulate the expression of POSH polypeptides or POSH-APs (e.g., HERPUD1). In certain embodiments, RNAi therapeutics are delivered to the liver (e.g., to hepatocytes). Exemplary RNAi therapeutics include any one of SEQ ID NOs: 15, 16, 18, 19, 21, 22, 24 and 25.

Therapeutic polypeptides may be generated by designing polypeptides to mimic certain protein domains important in the formation of POSH: POSH-AP complexes, such as, for example, SH3 or RING domains. For example, a polypeptide comprising a POSH SH3 domain such as, for example, the SH3 domain as set forth in SEQ ID NO: 30 will compete for binding to a POSH SH3 domain and will therefore act to disrupt binding of a partner protein. In one embodiment, a binding partner may be a Gag polypeptide. In another embodiment, a binding partner may be Rac. In a further embodiment, a polypeptide that resembles an L domain may disrupt recruitment of Gag to the POSH complex.

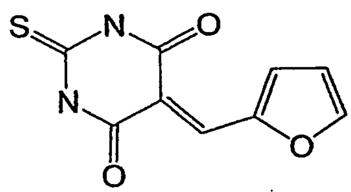
In view of the specification, methods for generating antibodies directed to epitopes of POSH and POSH-APs are known in the art. Antibodies may be introduced into cells by a variety of methods. One exemplary method comprises generating a nucleic acid encoding a single chain antibody that is capable of disrupting a POSH:POSH-AP complex. Such a nucleic acid may be conjugated to antibody that binds to receptors on the surface of target cells. It is contemplated that in certain embodiments, the antibody may target viral proteins that are present on the surface of infected cells, and in this way deliver the nucleic acid only to infected cells. Once bound to the target cell surface, the antibody is taken up by endocytosis, and the conjugated nucleic acid is transcribed and translated to produce a single chain antibody that interacts with and disrupts the targeted POSH:POSH-AP complex. Nucleic acids expressing the desired single chain antibody may also be

introduced into cells using a variety of more conventional techniques, such as viral transfection (e.g., using an adenoviral system) or liposome-mediated transfection.

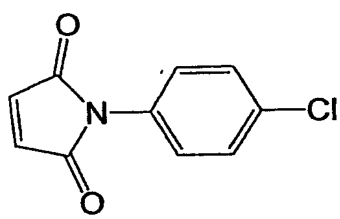
Small molecules of the application may be identified for their ability to modulate the formation of POSH:POSH-AP complexes.

Certain embodiments of the disclosure relate to use of a small molecule as an inhibitor of POSH. Examples of such small molecules include the following compounds:

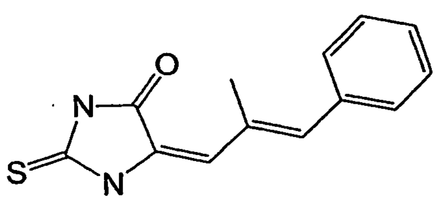
Compound CAS 27430-18-8:



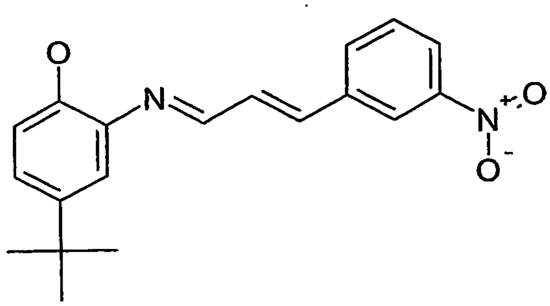
Compound CAS 1631-29-4:



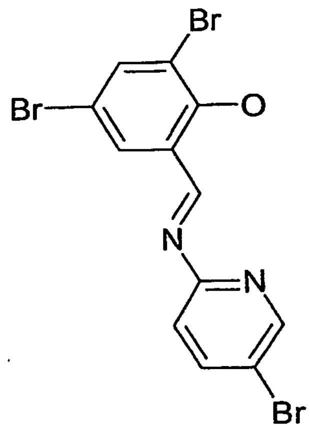
Compound CAS 503065-65-4:



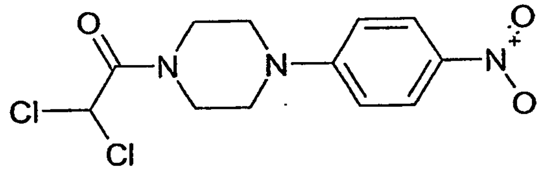
Compound CAS 414908-08:



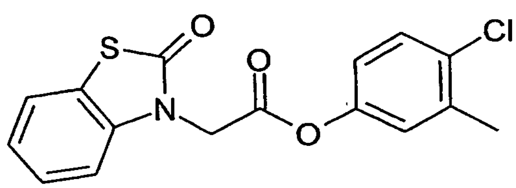
Compound CAS 415703-60-5:



Compound CAS 77367-94-3:



Compound CAS 154184-27-7:



In certain embodiments, compounds useful in the instant compositions and methods include heteroarylmethylene-dihydro-2,4,6-pyrimidinetriones and their thione analogs. Preferred heteroaryl moieties include 5-membered rings such as thienyl, furyl, pyrrolyl, oxazolyl, thiazolyl, and imidazolyl moieties.

In certain embodiments, compounds useful in the instant compositions and methods include N-arylmaleimides, especially N-phenylmaleimides, in which the phenyl group may be substituted or unsubstituted.

In certain embodiments, compounds useful in the instant compositions and methods include arylallylidene-2,4-imidazolidinediones and their thione analogs.

Preferred aryl groups are phenyl groups, and both the aryl and allylidene portions of the molecule may be substituted or unsubstituted.

In certain embodiments, compounds useful in the instant compositions and methods include substituted distyryl compounds and aza analogs thereof such as substituted 1,4-diphenylazabutadiene compounds.

In certain other embodiments, compounds useful in the instant compositions and methods include substituted styrenes and aza analogs thereof, such as 1,2-diphenylazaethylenes and 1-phenyl-2-pyridyl-azaethelenes.

In yet other embodiments, compounds useful in the instant compositions and methods include N-aryl-N'-acylpiperazines. In such compounds, the aryl ring, the acyl substituent, and/or the piperazine ring may be substituted or unsubstituted.

In additional embodiments, compounds useful in the instant compositions and methods include aryl esters of (2-oxo-benzooxazol-3-yl)-acetic acid, and analogs thereof in which one or more oxygen atoms are replaced by sulfur atoms.

In certain embodiments, the present application contemplates use of known PKA modulators (e.g., inhibitors or activators) in the methods of ihibiting viral infection and in the methods of treating or preventing cancer. Such PKA modulators include any compound, peptide, nucleotide derivative, nucleoside derivative, polysaccharide, sugar or other substance that can inhibit the activity of protein kinase A. Many PKA inhibitors are available and may be used. For example, many examples of PKA inhibitors including chemical structures, methods for administration and pharmacological effects are listed at the Calbiochem website at calbiochem.com. In general, inhibitors that also significantly inhibit protein kinase C activity are avoided.

In some embodiments, the PKA inhibitor is a nucleotide or nucleoside derivative. Specific examples of nucleoside or nucleotide derivatives that act as PKA inhibitors and that can be utilized in the disclosure include adenosine 3',5' cyclic monophosphorothioate. The H-89 inhibitor is a potent PKA inhibitor that can be used in the disclosure. The chemical name for the H-89 inhibitor is N-[2-((Pbromocinnamyl) amino)ethyl] isoquinolinesulfonamide. The KT5720 inhibitor from Calbiochem can also be used in the disclosure. Other PKA inhibitors which are available at from Calbiochem and can be used in the disclosure include ellagic acid

(also named 4,4',5,5',6,6'-hexahydroxydiphenic acid 2,6,2',6'-ditactone), piceatannol, 1-(5-Isoquinolinesulfonyl) methylpiperazine (H-7), N-[2-(methylamino)ethyl] isoquinolinesulfonamide (H-8), N-(2-aminoethyl) isoquinolinesulfonamide (H-9), and (5-isoquinolinesulfonyl)piperazine, 2HCI (H-100).

The PKA inhibitor can also be a peptide inhibitor (PKI). Such a peptide inhibitor can be any peptide that is recognized and bound by PKA but that PKA cannot phosphorylate. An example of a peptide inhibitor is a peptide with a "consensus sequence" for PKA recognition but with alanine in place of serine, for example, a peptide with the following sequence: Xaa-Arg-Arg-Xaa-Ala-Xaa, wherein Xaa is any amino acid, which specifically binds to the pseudoregion of the regulatory domain of PKA. Myristoylated PKA inhibitor amide (14-22, Cell-Permeable) having the sequence Myr-N-Gly-Arg-Thr-Gly-Arg-Arg-Asn-Ala-Ile-NH_2 is another example of a peptide inhibitor that can be utilized in the disclosure. A variety of other PKI peptides can be used as an inhibitor of protein kinase A in the practice of the disclosure. For example, several PKI peptides can be found in the NCBI protein database. See website at ncbi.nlm.nih.gov/ Genbank/GenbankOverview. One example of a human PKI peptide can be found at Genbank Accession No. P04541 (gi: 417194). Another example of a human PKI peptide is at Genbank Accession No. NP 008997 (gi: 5902020). Another PKI that can be used as an inhibitor has the following sequence: Ile-Ala-Ser-Gly-Arg-Thr-Gly-Arg-Arg-Asn-Ala-Ile-His-Asp-I1e-Leu-Val-SerSer-Ala. See published PCT application WO 03/080649.

Further examples of protein kinase A inhibitors are provided in the following references: Muniz et al., Proceedings of the National Academy of Sciences USA 1997 Dec 23; 94(26) 14461-66; Baude et al., Journal of Biological Chemistry Vol. 269 issue 27 18128-18133 (Jul. 1994); Scott et al.

Applicants found that POSH is phosphorylated by PKA and phosphorylation of POSH by PKA can inhibit POSH function, for example dissociating POSH from POSH interacting proteins (e.g, Rac). Therefore, in certain embodiments, the present disclosure also cotemplates use of PKA activators in treating or preventing a POSH-associated disease (e.g., viral infection or cancer). Exemplary PKA activators include, but are not limited to, forskolin, 8-Br-cAMP, and rolipram.

In additional embodiments of the application, compounds useful in the present application include phosphatase inhibitors. Phosphatase inhibitors useful in the subject application include sodium phosphate, sodium vanadate, and okadaic acid. In certain embodiments, the present application contemplates use of known phosphatase inhibitors in the methods of inhibiting viral infection, in the methods of treating or preventing cancer, and in the methods of inhibiting the progression of a neurodenerative disorder. Phosphatase inhibitors may be useful in inhibiting the activity of a POSH-AP, such as for example, PTPN12.

For POSH-APs that are GTPases, inhibitors such as GTPgamma35S would be effective at inhibiting the GTPase activity of the POSH-AP. For example, inhibition of ARF1 or ARF5 could be accomplished with the use of a GTPase inhibitor such as GTPgamma35S, a non-hydrolyzable form of GTP.

The generation of nucleic acid based therapeutic agents directed to POSH and POSH-APs is described below.

Methods for identifying and evaluating further modulators of POSH and POSH-APs are also provided below.

5. <u>RNA Interference, Ribozymes, Antisense and Related Constructs</u>

In certain aspects, the application relates to RNAi, ribozyme, antisense and other nucleic acid-related methods and compositions for manipulating (typically decreasing) a POSH activity. Exemplary RNAi and ribozyme molecules may comprise a sequence as shown in any of SEQ ID Nos: 15, 16, 18, 19, 21, 22, 24 and 25.

In certain aspects, the application relates to RNAi, ribozyme, antisense and other nucleic acid-related methods and compositions for manipulating (typically decreasing) a POSH-AP activity. Specific instances of nucleic acids that may be used to design nucleic acids for RNAi, ribozyme, antisense are provided in Figure 36. Additionally, nucleic acids of POSH-APs listed in Table 8 may be used to design nucleic acids for RNAi, ribozyme, antisense.

Certain embodiments of the application make use of materials and methods for effecting knockdown of one or more POSH or POSH-AP genes by means of RNA interference (RNAi). RNAi is a process of sequence-specific post-

transcriptional gene repression which can occur in eukaryotic cells. In general, this process involves degradation of an mRNA of a particular sequence induced by double-stranded RNA (dsRNA) that is homologous to that sequence. For example, the expression of a long dsRNA corresponding to the sequence of a particular single-stranded mRNA (ss mRNA) will labilize that message, thereby "interfering" with expression of the corresponding gene. Accordingly, any selected gene may be repressed by introducing a dsRNA which corresponds to all or a substantial part of the mRNA for that gene. It appears that when a long dsRNA is expressed, it is initially processed by a ribonuclease III into shorter dsRNA oligonucleotides of as few as 21 to 22 base pairs in length. Furthermore, Accordingly, RNAi may be effected by introduction or expression of relatively short homologous dsRNAs. Indeed the use of relatively short homologous dsRNAs may have certain advantages as discussed below.

Mammalian cells have at least two pathways that are affected by double-stranded RNA (dsRNA). In the RNAi (sequence-specific) pathway, the initiating dsRNA is first broken into short interfering (si) RNAs, as described above. The siRNAs have sense and antisense strands of about 21 nucleotides that form approximately 19 nucleotide si RNAs with overhangs of two nucleotides at each 3' end. Short interfering RNAs are thought to provide the sequence information that allows a specific messenger RNA to be targeted for degradation. In contrast, the nonspecific pathway is triggered by dsRNA of any sequence, as long as it is at least about 30 base pairs in length. The nonspecific effects occur because dsRNA activates two enzymes: PKR, which in its active form phosphorylates the translation initiation factor eIF2 to shut down all protein synthesis, and 2', 5' oligoadenylate synthetase (2', 5'-AS), which synthesizes a molecule that activates Rnase L, a nonspecific enzyme that targets all mRNAs. The nonspecific pathway may represent a host response to stress or viral infection, and, in general, the effects of the nonspecific pathway are preferably minimized under preferred methods of the present application. Significantly, longer dsRNAs appear to be required to induce the nonspecific pathway and, accordingly, dsRNAs shorter than about 30 bases pairs are preferred to effect gene repression by RNAi (see Hunter et al. (1975) J Biol

Chem 250: 409-17; Manche et al. (1992) Mol Cell Biol 12: 5239-48; Minks et al. (1979) J Biol Chem 254: 10180-3; and Elbashir et al. (2001) Nature 411: 494-8).

RNAi has been shown to be effective in reducing or eliminating the expression of genes in a number of different organisms including Caenorhabditiis elegans (see e.g., Fire et al. (1998) Nature 391: 806-11), mouse eggs and embryos (Wianny et al. (2000) Nature Cell Biol 2: 70-5; Svoboda et al. (2000) Development 127: 4147-56), and cultured RAT-1 fibroblasts (Bahramina et al. (1999) Mol Cell Biol 19: 274-83), and appears to be an anciently evolved pathway available in eukaryotic plants and animals (Sharp (2001) Genes Dev. 15: 485-90). RNAi has proven to be an effective means of decreasing gene expression in a variety of cell types including HeLa cells, NIH/3T3 cells, COS cells, 293 cells and BHK-21 cells, and typically decreases expression of a gene to lower levels than that achieved using antisense techniques and, indeed, frequently eliminates expression entirely (see Bass (2001) Nature 411: 428-9). In mammalian cells, siRNAs are effective at concentrations that are several orders of magnitude below the concentrations typically used in antisense experiments (Elbashir et al. (2001) Nature 411: 494-8).

The double stranded oligonucleotides used to effect RNAi are preferably less than 30 base pairs in length and, more preferably, comprise about 25, 24, 23, 22, 21, 20, 19, 18 or 17 base pairs of ribonucleic acid. Optionally the dsRNA oligonucleotides of the application may include 3' overhang ends. Exemplary 2-nucleotide 3' overhangs may be composed of ribonucleotide residues of any type and may even be composed of 2'-deoxythymidine resides, which lowers the cost of RNA synthesis and may enhance nuclease resistance of siRNAs in the cell culture medium and within transfected cells (see Elbashir et al. (2001) Nature 411: 494-8). Longer dsRNAs of 50, 75, 100 or even 500 base pairs or more may also be utilized in certain embodiments of the application. Exemplary concentrations of dsRNAs for effecting RNAi are about 0.05 nM, 0.1 nM, 0.5 nM, 1.0 nM, 1.5 nM, 25 nM or 100 nM, although other concentrations may be utilized depending upon the nature of the cells treated, the gene target and other factors readily discernable the skilled artisan. Exemplary dsRNAs may be synthesized chemically or produced in vitro or in vivo using appropriate expression vectors. Exemplary synthetic RNAs include 21 nucleotide RNAs chemically synthesized using methods known in the art (e.g.,

Expedite RNA phophoramidites and thymidine phosphoramidite (Proligo, Germany). Synthetic oligonucleotides are preferably deprotected and gel-purified using methods known in the art (see e.g., Elbashir et al. (2001) Genes Dev. 15: 188-200). Longer RNAs may be transcribed from promoters, such as T7 RNA polymerase promoters, known in the art. A single RNA target, placed in both possible orientations downstream of an in vitro promoter, will transcribe both strands of the target to create a dsRNA oligonucleotide of the desired target sequence. Any of the above RNA species will be designed to include a portion of nucleic acid sequence represented in a POSH or POSH-AP nucleic acid, such as, for example, a nucleic acid that hybridizes, under stringent and/or physiological conditions, to any of SEQ ID Nos: 1, 3, 4, 6, 8 and 10 and complements thereof or any of the POSH-AP sequences presented in Figure 36.

The specific sequence utilized in design of the oligonucleotides may be any contiguous sequence of nucleotides contained within the expressed gene message of the target. Programs and algorithms, known in the art, may be used to select appropriate target sequences. In addition, optimal sequences may be selected utilizing programs designed to predict the secondary structure of a specified single stranded nucleic acid sequence and allowing selection of those sequences likely to occur in exposed single stranded regions of a folded mRNA. Methods and compositions for designing appropriate oligonucleotides may be found, for example, in U.S. Patent Nos. 6,251,588, the contents of which are incorporated herein by reference. Messenger RNA (mRNA) is generally thought of as a linear molecule which contains the information for directing protein synthesis within the sequence of ribonucleotides, however studies have revealed a number of secondary and tertiary structures that exist in most mRNAs. Secondary structure elements in RNA are formed largely by Watson-Crick type interactions between different regions of the same RNA molecule. Important secondary structural elements include intramolecular double stranded regions, hairpin loops, bulges in duplex RNA and internal loops. Tertiary structural elements are formed when secondary structural elements come in contact with each other or with single stranded regions to produce a more complex three dimensional structure. A number of researchers have measured the binding energies of a large number of RNA duplex structures and have

derived a set of rules which can be used to predict the secondary structure of RNA (see e.g., Jaeger et al. (1989) Proc. Natl. Acad. Sci. USA 86:7706 (1989); and Turner et al. (1988) Annu. Rev. Biophys. Biophys. Chem. 17:167) . The rules are useful in identification of RNA structural elements and, in particular, for identifying single stranded RNA regions which may represent preferred segments of the mRNA to target for silencing RNAi, ribozyme or antisense technologies. Accordingly, preferred segments of the mRNA target can be identified for design of the RNAi mediating dsRNA oligonucleotides as well as for design of appropriate ribozyme and hammerheadribozyme compositions of the application.

The dsRNA oligonucleotides may be introduced into the cell by transfection with an heterologous target gene using carrier compositions such as liposomes, which are known in the art- e.g., Lipofectamine 2000 (Life Technologies) as described by the manufacturer for adherent cell lines. Transfection of dsRNA oligonucleotides for targeting endogenous genes may be carried out using Oligofectamine (Life Technologies). Transfection efficiency may be checked using fluorescence microscopy for mammalian cell lines after co-transfection of hGFP-encoding pAD3 (Kehlenback et al. (1998) J Cell Biol 141: 863-74). The effectiveness of the RNAi may be assessed by any of a number of assays following introduction of the dsRNAs. These include Western blot analysis using antibodies which recognize the POSH or POSH-AP gene product following sufficient time for turnover of the endogenous pool after new protein synthesis is repressed, reverse transcriptase polymerase chain reaction and Northern blot analysis to determine the level of existing POSH or POSH-AP target mRNA.

Further compositions, methods and applications of RNAi technology are provided in U.S. Patent Application Nos. 6,278,039, 5,723,750 and 5,244,805, which are incorporated herein by reference.

Ribozyme molecules designed to catalytically cleave POSH or POSH-AP mRNA transcripts can also be used to prevent translation of suject POSH or POSH-AP mRNAs and/or expression of POSH or POSH-APs (see, e.g., PCT International Publication WO90/11364, published October 4, 1990; Sarver et al. (1990) Science 247:1222-1225 and U.S. Patent No. 5,093,246). Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. (For a review, see

Rossi (1994) Current Biology 4: 469-471). The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by an endonucleolytic cleavage event. The composition of ribozyme molecules preferably includes one or more sequences complementary to a POSH or POSH-AP mRNA, and the well known catalytic sequence responsible for mRNA cleavage or a functionally equivalent sequence (see, e.g., U.S. Pat. No. 5,093,246, which is incorporated herein by reference in its entirety).

While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy target mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. Preferably, the target mRNA has the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach ((1988) Nature 334:585-591; and see PCT Appln. No. WO89/05852, the contents of which are incorporated herein by reference). Hammerhead ribozyme sequences can be embedded in a stable RNA such as a transfer RNA (tRNA) to increase cleavage efficiency in vivo (Perriman et al. (1995) Proc. Natl. Acad. Sci. USA, 92: 6175-79; de Feyter, and Gaudron, Methods in Molecular Biology, Vol. 74, Chapter 43, "Expressing Ribozymes in Plants", Edited by Turner, P. C, Humana Press Inc., Totowa, N.J). In particular, RNA polymerase III-mediated expression of tRNA fusion ribozymes are well known in the art (see Kawasaki et al. (1998) Nature 393: 284-9; Kuwabara et al. (1998) Nature Biotechnol. 16: 961-5; and Kuwabara et al. (1998) Mol. Cell 2: 617-27; Koseki et al. (1999) J Virol 73: 1868-77; Kuwabara et al. (1999) Proc Natl Acad Sci USA 96: 1886-91; Tanabe et al. (2000) Nature 406: 473-4). There are typically a number of potential hammerhead ribozyme cleavage sites within a given target cDNA sequence. Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the target mRNA- to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts. Furthermore, the use of any cleavage recognition site located in the target sequence encoding different portions of the C-terminal amino acid domains of, for example,

long and short forms of target would allow the selective targeting of one or the other form of the target, and thus, have a selective effect on one form of the target gene product.

Gene targeting ribozymes necessarily contain a hybridizing region complementary to two regions, each of at least 5 and preferably each 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 contiguous nucleotides in length of a POSH or POSH-AP mRNA, such as an mRNA of a sequence represented in any of SEQ ID Nos: 1, 3, 4, 6, 8 or 10 or a POSH-AP presented in Figure 36. In addition, ribozymes possess highly specific endoribonuclease activity, which autocatalytically cleaves the target sense mRNA. The present application extends to ribozymes which hybridize to a sense mRNA encoding a POSH gene such as a therapeutic drug target candidate gene, thereby hybridising to the sense mRNA and cleaving it, such that it is no longer capable of being translated to synthesize a functional polypeptide product.

The ribozymes of the present application also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in Tetrahymena thermophila (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug, et al. (1984) Science 224:574-578; Zaug, et al. (1986) Science 231:470-475; Zaug, et al. (1986) Nature 324:429-433; published International patent application No. WO88/04300 by University Patents Inc.; Been, et al. (1986) Cell 47:207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The application encompasses those Cech-type ribozymes which target eight base-pair active site sequences that are present in a target gene or nucleic acid sequence.

Ribozymes can be composed of modified oligonucleotides (e.g., for improved stability, targeting, etc.) and should be delivered to cells which express the target gene in vivo. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous target messages and inhibit translation. Because ribozymes,

unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

In certain embodiments, a ribozyme may be designed by first identifying a sequence portion sufficient to cause effective knockdown by RNAi. The same sequence portion may then be incorporated into a ribozyme. In this aspect of the application, the gene-targeting portions of the ribozyme or RNAi are substantially the same sequence of at least 5 and preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more contiguous nucleotides of a POSH nucleic acid, such as a nucleic acid of any of SEQ ID Nos: 1, 3, 4, 6, 8, or 10 or POSH-AP nucleic acid, as presented in Figure 36. In a long target RNA chain, significant numbers of target sites are not accessible to the ribozyme because they are hidden within secondary or tertiary structures (Birikh et al. (1997) Eur J Biochem 245: 1-16). To overcome the problem of target RNA accessibility, computer generated predictions of secondary structure are typically used to identify targets that are most likely to be single-stranded or have an "open" configuration (see Jaeger et al. (1989) Methods Enzymol 183: 281-306). Other approaches utilize a systematic approach to predicting secondary structure which involves assessing a huge number of candidate hybridizing oligonucleotides molecules (seeMilner et al. (1997) Nat Biotechnol 15: 537-41; and Patzel and Sczakiel (1998) Nat Biotechnol 16: 64-8). Additionally, U.S. Patent No. 6,251,588, the contents of which are hereby incorporated herein, describes methods for evaluating oligonucleotide probe sequences so as to predict the potential for hybridization to a target nucleic acid sequence. The method of the application provides for the use of such methods to select preferred segments of a target mRNA sequence that are predicted to be single-stranded and, further, for the opportunistic utilization of the same or substantially identical target mRNA sequence, preferably comprising about 10-20 consecutive nucleotides of the target mRNA, in the design of both the RNAi oligonucleotides and ribozymes of the application.

A further aspect of the application relates to the use of the isolated "antisense" nucleic acids to inhibit expression, e.g., by inhibiting transcription and/or translation of a POSH or POSH-AP nucleic acid. The antisense nucleic acids may bind to the potential drug target by conventional base pair complementarity, or,

for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix. In general, these methods refer to the range of techniques generally employed in the art, and include any methods that rely on specific binding to oligonucleotide sequences.

An antisense construct of the present application can be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of the cellular mRNA which encodes a POSH or POSH-AP polypeptide. Alternatively, the antisense construct is an oligonucleotide probe, which is generated ex vivo and which, when introduced into the cell causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences of a POSH or POSH-AP nucleic acid. Such oligonucleotide probes are preferably modified oligonucleotides, which are resistant to endogenous nucleases, e.g., exonucleases and/or endonucleases, and are therefore stable in vivo. Exemplary nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patents 5,176,996; 5,264,564; and 5,256,775). Additionally, general approaches to constructing oligomers useful in antisense therapy have been reviewed, for example, by Van der Krol et al. (1988) BioTechniques 6:958-976; and Stein et al. (1988) Cancer Res 48:2659- 2668.

With respect to antisense DNA, oligodeoxyribonucleotides derived from the translation initiation site, e.g., between the -10 and +10 regions of the target gene, are preferred. Antisense approaches involve the design of oligonucleotides (either DNA or RNA) that are complementary to mRNA encoding a POSH or POSH-AP polypeptide. The antisense oligonucleotides will bind to the mRNA transcripts and prevent translation. Absolute complementarity, although preferred, is not required. In the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of

mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the mRNA, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have recently been shown to be effective at inhibiting translation of mRNAs as well. (Wagner, R. 1994. Nature 372:333). Therefore, oligonucleotides complementary to either the 5' or 3' untranslated, non-coding regions of a gene could be used in an antisense approach to inhibit translation of that mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could also be used in accordance with the application. Whether designed to hybridize to the 5', 3' or coding region of mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably less that about 100 and more preferably less than about 50, 25, 17 or 10 nucleotides in length.

It is preferred that in vitro studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an internal control RNA or protein. Results obtained using the antisense oligonucleotide may be compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

The antisense oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for

targeting host cell receptors), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810, published December 15, 1988) or the blood- brain barrier (see, e.g., PCT Publication No. W089/10134, published April 25, 1988), hybridization-triggered cleavage agents. (See, e.g., Krol et al., 1988, BioTechniques 6:958- 976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5- bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5- (carboxyhydroxytiethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5- carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6- isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6- isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5- oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3- N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

The antisense oligonucleotide can also contain a neutral peptide-like backbone. Such molecules are termed peptide nucleic acid (PNA)-oligomers and are described, e.g., in Perry-O'Keefe et al. (1996) Proc. Natl. Acad. Sci. U.S.A. 93:14670 and in Eglom et al. (1993) Nature 365:566. One advantage of PNA

oligomers is their capability to bind to complementary DNA essentially independently from the ionic strength of the medium due to the neutral backbone of the DNA. In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet a further embodiment, the antisense oligonucleotide is an alpha-anomeric oligonucleotide. An alpha-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual antiparallel orientation, the strands run parallel to each other (Gautier et al., 1987, Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-0-methylribonucleotide (Inoue et al., 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue et al., 1987, FEBS Lett. 215:327-330).

While antisense nucleotides complementary to the coding region of a POSH or POSH-AP mRNA sequence can be used, those complementary to the transcribed untranslated region may also be used.

In certain instances, it may be difficult to achieve intracellular concentrations of the antisense sufficient to suppress translation on endogenous mRNAs. Therefore a preferred approach utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. The use of such a construct to transfect target cells will result in the transcription of sufficient amounts of single stranded RNAs that will form complementary base pairs with the endogenous potential drug target transcripts and thereby prevent translation. For example, a vector can be introduced such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can

be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster et al, 1982, Nature 296:39-42), etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct, which can be introduced directly into the tissue site.

Alternatively, POSH or POSH-AP gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the gene (i.e., the promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells in the body. (See generally, Helene, C. 1991, Anticancer Drug Des., 6(6):569-84; Helene, C., et al., 1992, Ann. N.Y. Acad. Sci., 660:27-36; and Maher, L.J., 1992, Bioassays 14(12):807-15).

Nucleic acid molecules to be used in triple helix formation for the inhibition of transcription are preferably single stranded and composed of deoxyribonucleotides. The base composition of these oligonucleotides should promote triple helix formation via Hoogsteen base pairing rules, which generally require sizable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which will result in TAT and CGC triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine- rich, for example, containing a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in CGC triplets across the three strands in the triplex.

Alternatively, POSH or POSH-AP sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other,

eliminating the necessity for a sizable stretch of either purines or pyrimidines to be present on one strand of a duplex.

A further aspect of the application relates to the use of DNA enzymes to inhibit expression of a POSH or POSH-AP gene. DNA enzymes incorporate some of the mechanistic features of both antisense and ribozyme technologies. DNA enzymes are designed so that they recognize a particular target nucleic acid sequence, much like an antisense oligonucleotide, however much like a ribozyme they are catalytic and specifically cleave the target nucleic acid.

There are currently two basic types of DNA enzymes, and both of these were identified by Santoro and Joyce (see, for example, US Patent No. 6110462). The 10-23 DNA enzyme comprises a loop structure which connect two arms. The two arms provide specificity by recognizing the particular target nucleic acid sequence while the loop structure provides catalytic function under physiological conditions.

Briefly, to design an ideal DNA enzyme that specifically recognizes and cleaves a target nucleic acid, one of skill in the art must first identify the unique target sequence. This can be done using the same approach as outlined for antisense oligonucleotides. Preferably, the unique or substantially sequence is a G/C rich of approximately 18 to 22 nucleotides. High G/C content helps insure a stronger interaction between the DNA enzyme and the target sequence.

When synthesizing the DNA enzyme, the specific antisense recognition sequence that will target the enzyme to the message is divided so that it comprises the two arms of the DNA enzyme, and the DNA enzyme loop is placed between the two specific arms.

Methods of making and administering DNA enzymes can be found, for example, in US 6110462. Similarly, methods of delivery DNA ribozymes in vitro or in vivo include methods of delivery RNA ribozyme, as outlined in detail above. Additionally, one of skill in the art will recognize that, like antisense oligonucleotide, DNA enzymes can be optionally modified to improve stability and improve resistance to degradation.

Antisense RNA and DNA, ribozyme, RNAi and triple helix molecules of the application may be prepared by any method known in the art for the synthesis of DNA and RNA molecules. These include techniques for chemically synthesizing

oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines. Moreover, various well-known modifications to nucleic acid molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to the addition of flanking sequences of ribonucleotides or deoxyribonucleotides to the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

6. <u>Drug Screening Assays</u>

In certain aspects, the present application provides assays for identifying therapeutic agents which either interfere with or promote POSH or POSH-AP function. In certain aspects, the present application also provides assays for identifying therapeutic agents which either interfere with or promote the complex formation between a POSH polypeptide and a POSH-AP polypeptide.

In certain embodiments, agents of the application are antiviral agents, optionally interfering with viral maturation, and preferably where the virus is an envelope virus, and optionally a retroid virus or an RNA virus. In other embodiments, agents of the application are anticancer agents. In further embodiments, agents of the application inhibit the progression of a neurodegenerative disorder. In certain embodiments, an antiviral or anticancer agent or an agent that inhibits the progression of a neurodegenerative disorder interferes with the ubiquitin ligase catalytic activity of POSH (e.g., POSH auto-ubiquitination or transfer to a target protein). In other embodiments, agents disclosed herein inhibit or promote POSH and POSH-AP mediated cellular processes such as apoptosis and protein processing in the secretory pathway.

In certain preferred embodiments, an antiviral agent interferes with the interaction between POSH and a POSH-AP polypeptide, for example an antiviral agent may disrupt or render irreversible interaction between a POSH polypeptide and POSH-AP polypeptide (as in the case of a POSH dimer, a heterodimer of two different POSH polypeptides, homomultimers and heteromultimers). In further embodiments, agents of the application are anti-apoptotic agents, optionally interfering with JNK and/or NF-κB signaling. In yet additional embodiments, agents of the application interfere with the signaling of a GTPase, such as Rac or Ras, optionally disrupting the interaction between a POSH polypeptide and a Rac protein. In certain embodiments, agents of the application modulate the ubiquitin ligase activity of POSH and may be used to treat certain diseases related to ubiquitin ligase activity. In certain embodiments, agents of the application interfere with the trafficking of a protein through the secretory pathway.

In certain embodiments, the application provides assays to identify, optimize or otherwise assess agents that increase or decrease a ubiquitin-related activity of a POSH polypeptide. Ubiquitin-related activities of POSH polypeptides may include the self-ubiquitination activity of a POSH polypeptide, generally involving the transfer of ubiquitin from an E2 enzyme to the POSH polypeptide, and the ubiquitination of a target protein, generally involving the transfer of a ubiquitin from a POSH polypeptide to the target protein. In certain embodiments, a POSH activity is mediated, at least in part, by a POSH RING domain.

In certain embodiments, an assay comprises forming a mixture comprising a POSH polypeptide, an E2 polypeptide and a source of ubiquitin (which may be the E2 polypeptide pre-complexed with ubiquitin). Optionally the mixture comprises an E1 polypeptide and optionally the mixture comprises a target polypeptide. Additional components of the mixture may be selected to provide conditions consistent with the ubiquitination of the POSH polypeptide. One or more of a variety of parameters may be detected, such as POSH-ubiquitin conjugates, E2-ubiquitin thioesters, free ubiquitin and target polypeptide-ubiquitin complexes. The term "detect" is used herein to include a determination of the presence or absence of the subject of detection (e.g., POSH-ubiqutin, E2-ubiquitin, etc.), a quantitative measure of the amount of the subject of detection, or a mathematical calculation of

the presence, absence or amount of the subject of detection, based on the detection of other parameters. The term "detect" includes the situation wherein the subject of detection is determined to be absent or below the level of sensitivity. Detection may comprise detection of a label (e.g., fluorescent label, radioisotope label, and other described below), resolution and identification by size (e.g., SDS-PAGE, mass spectroscopy), purification and detection, and other methods that, in view of this specification, will be available to one of skill in the art. For instance, radioisotope labeling may be measured by scintillation counting, or by densitometry after exposure to a photographic emulsion, or by using a device such as a Phosphorimager. Likewise, densitometry may be used to measure bound ubiquitin following a reaction with an enzyme label substrate that produces an opaque product when an enzyme label is used. In a preferred embodiment, an assay comprises detecting the POSH-ubiquitin conjugate.

In certain embodiments, an assay comprises forming a mixture comprising a POSH polypeptide, a target polypeptide and a source of ubiquitin (which may be the POSH polypeptide pre-complexed with ubiquitin). Optionally the mixture comprises an E1 and/or E2 polypeptide and optionally the mixture comprises an E2-ubiquitin thioester. Additional components of the mixture may be selected to provide conditions consistent with the ubiquitination of the target polypeptide. One or more of a variety of parameters may be detected, such as POSH-ubiquitin conjugates and target polypeptide-ubiquitin conjugates. In a preferred embodiment, an assay comprises detecting the target polypeptide-ubiquitin conjugate. In another preferred embodiment, an assay comprises detecting the POSH-ubiquitin conjugate.

An assay described above may be used in a screening assay to identify agents that modulate a ubiquitin-related activity of a POSH polypeptide. A screening assay will generally involve adding a test agent to one of the above assays, or any other assay designed to assess a ubiquitin-related activity of a POSH polypeptide. The parameter(s) detected in a screening assay may be compared to a suitable reference. A suitable reference may be an assay run previously, in parallel or later that omits the test agent. A suitable reference may also be an average of previous measurements in the absence of the test agent. In general the components of a screening assay mixture may be added in any order consistent with the overall

activity to be assessed, but certain variations may be preferred. For example, in certain embodiments, it may be desirable to pre-incubate the test agent and the E3 (e.g., the POSH polypeptide), followed by removing the test agent and addition of other components to complete the assay. In this manner, the effects of the agent solely on the POSH polypeptide may be assessed. In certain preferred embodiments, a screening assay for an antiviral agent employs a target polypeptide comprising an L domain, and preferably an HIV L domain.

In certain embodiments, an assay is performed in a high-throughput format. For example, one of the components of a mixture may be affixed to a solid substrate and one or more of the other components is labeled. For example, the POSH polypeptide may be affixed to a surface, such as a 96-well plate, and the ubiquitin is in solution and labeled. An E2 and E1 are also in solution, and the POSH-ubiquitin conjugate formation may be measured by washing the solid surface to remove uncomplexed labeled ubiquitin and detecting the ubiquitin that remains bound. Other variations may be used. For example, the amount of ubiquitin in solution may be detected. In certain embodiments, the formation of ubiquitin complexes may be measured by an interactive technique, such as FRET, wherein a ubiquitin is labeled with a first label and the desired complex partner (e.g., POSH polypeptide or target polypeptide) is labeled with a second label, wherein the first and second label interact when they come into close proximity to produce an altered signal. In FRET, the first and second labels are fluorophores. FRET is described in greater detail below. The formation of polyubiquitin complexes may be performed by mixing two or more pools of differentially labeled ubiquitin that interact upon formation of a polyubiqutin (see, e.g., US Patent Publication 20020042083). High-throughput may be achieved by performing an interactive assay, such as FRET, in solution as well. In addition, if a polypeptide in the mixture, such as the POSH polypeptide or target polypeptide, is readily purifiable (e.g., with a specific antibody or via a tag such as biotin, FLAG, polyhistidine, etc.), the reaction may be performed in solution and the tagged polypeptide rapidly isolated, along with any polypeptides, such as ubiquitin, that are associated with the tagged polypeptide. Proteins may also be resolved by SDS-PAGE for detection.

In certain embodiments, the ubiquitin is labeled, either directly or indirectly. This typically allows for easy and rapid detection and measurement of ligated ubiquitin, making the assay useful for high-throughput screening applications. As descrived above, certain embodiments may employ one or more tagged or labeled proteins. A "tag" is meant to include moieties that facilitate rapid isolation of the tagged polypeptide. A tag may be used to facilitate attachment of a polypeptide to a surface. A "label" is meant to include moieties that facilitate rapid detection of the labeled polypeptide. Certain moieties may be used both as a label and a tag (e.g., epitope tags that are readily purified and detected with a well-characterized antibody). Biotinylation of polypeptides is well known, for example, a large number of biotinylation agents are known, including amine-reactive and thiol-reactive agents, for the biotinylation of proteins, nucleic acids, carbohydrates, carboxylic acids; see chapter 4, Molecular Probes Catalog, Haugland, 6th Ed. 1996, hereby incorporated by reference. A biotinylated substrate can be attached to a biotinylated component via avidin or streptavidin. Similarly, a large number of haptenylation reagents are also known.

An "E1" is a ubiquitin activating enzyme. In a preferred embodiment, E1 is capable of transferring ubiquitin to an E2. In a preferred embodiment, E1 forms a high energy thiolester bond with ubiquitin, thereby "activating" the ubiquitin. An "E2" is a ubiquitin carrier enzyme (also known as a ubiquitin conjugating enzyme). In a preferred embodiment, ubiquitin is transferred from E1 to E2. In a preferred embodiment, the transfer results in a thiolester bond formed between E2 and ubiquitin. In a preferred embodiment, E2 is capable of transferring ubiquitin to a POSH polypeptide.

In an alternative embodiment, a POSH polypeptide, E2 or target polypeptide is bound to a bead, optionally with the assistance of a tag. Following ligation, the beads may be separated from the unbound ubiquitin and the bound ubiquitin measured. In a preferred embodiment, POSH polypeptide is bound to beads and the composition used includes labeled ubiquitin. In this embodiment, the beads with bound ubiquitin may be separated using a fluorescence-activated cell sorting (FACS) machine. Methods for such use are described in U.S. patent application Ser.

No. 09/047,119, which is hereby incorporated in its entirety. The amount of bound ubiquitin can then be measured.

In a screening assay, the effect of a test agent may be assessed by, for example, assessing the effect of the test agent on kinetics, steady-state and/or endpoint of the reaction.

The components of the various assay mixtures provided herein may be combined in varying amounts. In a preferred embodiment, ubiquitin (or E2 complexed ubiquitin) is combined at a final concentration of from 5 to 200 ng per 100 microliter reaction solution. Optionally E1 is used at a final concentration of from 1 to 50 ng per 100 microliter reaction solution. Optionally E2 is combined at a final concentration of 10 to 100 ng per 100 microliter reaction solution, more preferably 10-50 ng per 100 microliter reaction solution. In a preferred embodiment, POSH polypeptide is combined at a final concentration of from 1 to 500 ng per 100 microliter reaction solution.

Generally, an assay mixture is prepared so as to favor ubiquitin ligase activity and/or ubiquitination activity. Generally, this will be physiological conditions, such as 50 – 200 mM salt (e.g., NaCl, KCl), pH of between 5 and 9, and preferably between 6 and 8. Such conditions may be optimized through trial and error. Incubations may be performed at any temperature which facilitates optimal activity, typically between 4 and 40 °C. Incubation periods are selected for optimum activity, but may also be optimized to facilitate rapid high through put screening. Typically between 0.5 and 1.5 hours will be sufficient. A variety of other reagents may be included in the compositions. These include reagents like salts, solvents, buffers, neutral proteins, e.g., albumin, detergents, etc. which may be used to facilitate optimal ubiquitination enzyme activity and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The compositions will also preferably include adenosine tri-phosphate (ATP). The mixture of components may be added in any order that promotes ubiquitin ligase activity or optimizes identification of candidate modulator effects. In a preferred embodiment, ubiquitin is provided in a reaction buffer solution, followed by addition of the ubiquitination enzymes. In an alternate preferred embodiment,

ubiquitin is provided in a reaction buffer solution, a candidate modulator is then added, followed by addition of the ubiquitination enzymes.

In general, a test agent that decreases a POSH ubiquitin-related activity may be used to inhibit POSH function in vivo, while a test agent that increases a POSH ubiquitin-related activity may be used to stimulate POSH function in vivo. Test agent may be modified for use in vivo, e.g., by addition of a hydrophobic moiety, such as an ester.

In certain embodiments, a ubiquitination assay as described above for POSH can similarly be conducted for a Cbl-b, a SIAH1, or a TTC3 polypeptide. In certain embodiments, the application provides assays to identify, optimize or otherwise assess agents that increase or decrease a ubiquitin-related activity of a Cbl-b, a SIAH1, or a TTC3 polypeptide. Ubiquitin-related activities of Cbl-b, SIAH1, or TTC3 polypeptides may include the self-ubiquitination activity of a Cbl-b, SIAH1, or TTC3 polypeptide, generally involving the transfer of ubiquitin from an E2 enzyme to the respective Cbl-b, SIAH1, or TTC3 polypeptide, and the ubiquitination of a target protein, e.g., the p85 subunit of PI3K, e.g, synaptophysin, generally involving the transfer of a ubiquitin from a Cbl-b, SIAH1, or TTC3 polypeptide to the target protein, e.g, the p85 subunit of PI3K, e.g., synaptophysin, e.g., HERPUD1. In certain embodiments, a Cbl-b, a SIAH1, or a TTC3 activity is mediated, at least in part, by a RING domain of a Cbl-b, a SIAH1, or a TTC3, respectively.

An additional POSH-AP may be added to a POSH ubiquitination assay to assess the effect of the POSH-AP (e.g., PRKAR1A, PRKACA, or PRKACB) on POSH-mediated ubiquitination and/or to assess whether the POSH-AP is a target for POSH-mediated ubiquitination (e.g., HERPUD1, e.g., PKA).

Certain embodiments of the application relate to assays for identifying agents that bind to a POSH or POSH-AP polypeptide, optionally a particular domain of POSH such as an SH3 or RING domain or a particular domain of a POSH-AP, particularly a kinase catalytic domain or ATP binding domain. In preferred embodiments, a POSH polypeptide is a polypeptide comprising the fourth SH3 domain of hPOSH (SEQ ID NO: 30). A wide variety of assays may be used for this purpose, including labeled in vitro protein-protein binding assays, electrophoretic

mobility shift assays, immunoassays for protein binding, and the like. The purified protein may also be used for determination of three-dimensional crystal structure, which can be used for modeling intermolecular interactions and design of test agents. In one embodiment, an assay detects agents which inhibit interaction of one or more subject POSH polypeptides with a POSH-AP. In another embodiment, the assay detects agents which modulate the intrinsic biological activity of a POSH polypeptide or POSH complex, such as an enzymatic activity, binding to other cellular components, cellular compartmentalization, and the like.

Certain embodiments of the application relate to assays for identifying agents that modulate a POSH-AP polypeptide such as a PKA subunit polypeptide. Preferred PKA subunit polypeptides include PRKAR1A, PRKACA, and PRKACB. Exemplary assays used for this purpose may include detecting phosphorylation of PKA subunit, kinase activity of the PKA subunit, ability of the PKA subunit to elicit downstream signaling of the PKA pathway, and the like. For example, activity of protein kinase A can be assayed either in vitro or in vivo. PKA activity can be determined by detecting posphorylation of a PKA specific substrate. The specific PKA substrate can be any convenient peptide with a serine that is recognized as a phosphorylation site by PKA. For example, the peptide substrate can have the sequence: Leu-Arg-Arg-Ala- Ser-Leu-Gly.

In one aspect, the application provides methods and compositions for the identification of compositions that interfere with the function of POSH or POSH-AP polypeptides. Given the role of POSH polypeptides in viral production, compositions that perturb the formation or stability of the protein-protein interactions between POSH polypeptides and the proteins that they interact with, such as POSH-APs, and particularly POSH complexes comprising a viral protein, are candidate pharmaceuticals for the treatment of viral infections.

While not wishing to be bound to mechanism, it is postulated that POSH polypeptides promote the assembly of protein complexes that are important in release of virions and other biological processes. Complexes of the application may include a combination of a POSH polypeptide and a POSH-AP. Exemplary complexes may comprise one or more of the following: a POSH polypeptide (as in

the case of a POSH dimer, a heterodimer of two different POSH, homomultimers and heteromultimers); a HERPUD1 polypeptide; or an MSTP028 polypeptide.

In an assay for an antiviral or antiapoptotic agent, the test agent is assessed for its ability to disrupt or inhibit the formation of a complex of a POSH polypeptide and a small GTPase, such as a Rac polypeptide, particularly a human Rac polypeptide, such as Rac1.

A variety of assay formats will suffice and, in light of the present disclosure, those not expressly described herein will nevertheless be comprehended by one of ordinary skill in the art. Assay formats which approximate such conditions as formation of protein complexes, enzymatic activity, and even a POSH polypeptide-mediated membrane reorganization or vesicle formation activity, may be generated in many different forms, and include assays based on cell-free systems, e.g., purified proteins or cell lysates, as well as cell-based assays which utilize intact cells. Simple binding assays can also be used to detect agents which bind to POSH. Such binding assays may also identify agents that act by disrupting the interaction between a POSH polypeptide and a POSH interacting protein, or the binding of a POSH polypeptide or complex to a substrate. Agents to be tested can be produced, for example, by bacteria, yeast or other organisms (e.g., natural products), produced chemically (e.g., small molecules, including peptidomimetics), or produced recombinantly. In a preferred embodiment, the test agent is a small organic molecule, e.g., other than a peptide or oligonucleotide, having a molecular weight of less than about 2,000 daltons.

In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays of the present application which are performed in cell-free systems, such as may be developed with purified or semi-purified proteins or with lysates, are often preferred as "primary" screens in that they can be generated to permit rapid development and relatively easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the in vitro system, the assay instead being focused primarily on the effect of the drug on the molecular target as may be

manifest in an alteration of binding affinity with other proteins or changes in enzymatic properties of the molecular target.

In preferred in vitro embodiments of the present assay, a reconstituted POSH complex comprises a reconstituted mixture of at least semi-purified proteins. By semi-purified, it is meant that the proteins utilized in the reconstituted mixture have been previously separated from other cellular or viral proteins. For instance, in contrast to cell lysates, the proteins involved in POSH complex formation are present in the mixture to at least 50% purity relative to all other proteins in the mixture, and more preferably are present at 90-95% purity. In certain embodiments of the subject method, the reconstituted protein mixture is derived by mixing highly purified proteins such that the reconstituted mixture substantially lacks other proteins (such as of cellular or viral origin) which might interfere with or otherwise alter the ability to measure POSH complex assembly and/or disassembly.

Assaying POSH complexes, in the presence and absence of a candidate inhibitor, can be accomplished in any vessel suitable for containing the reactants. Examples include microtitre plates, test tubes, and micro-centrifuge tubes.

In one embodiment of the present application, drug screening assays can be generated which detect inhibitory agents on the basis of their ability to interfere with assembly or stability of the POSH complex. In an exemplary binding assay, the compound of interest is contacted with a mixture comprising a POSH polypeptide and at least one interacting polypeptide. Detection and quantification of POSH complexes provides a means for determining the compound's efficacy at inhibiting (or potentiating) interaction between the two polypeptides. The efficacy of the compound can be assessed by generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. In the control assay, the formation of complexes is quantitated in the absence of the test compound.

Complex formation between the POSH polypeptides and a substrate polypeptide may be detected by a variety of techniques, many of which are effectively described above. For instance, modulation in the formation of complexes can be quantitated using, for example, detectably labeled proteins (e.g., radiolabeled, fluorescently labeled, or enzymatically labeled), by immunoassay, or by

chromatographic detection. Surface plasmon resonance systems, such as those available from Biacore International AB (Uppsala, Sweden), may also be used to detect protein-protein interaction

Often, it will be desirable to immobilize one of the polypeptides to facilitate separation of complexes from uncomplexed forms of one of the proteins, as well as to accommodate automation of the assay. In an illustrative embodiment, a fusion protein can be provided which adds a domain that permits the protein to be bound to an insoluble matrix. For example, GST-POSH fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with a potential interacting protein, e.g., an ^{35}S-labeled polypeptide, and the test compound and incubated under conditions conducive to complex formation. Following incubation, the beads are washed to remove any unbound interacting protein, and the matrix bead-bound radiolabel determined directly (e.g., beads placed in scintillant), or in the supernatant after the complexes are dissociated, e.g., when microtitre plate is used. Alternatively, after washing away unbound protein, the complexes can be dissociated from the matrix, separated by SDS-PAGE gel, and the level of interacting polypeptide found in the matrix-bound fraction quantitated from the gel using standard electrophoretic techniques.

In a further embodiment, agents that bind to a POSH or POSH-AP may be identified by using an immobilized POSH or POSH-AP. In an illustrative embodiment, a fusion protein can be provided which adds a domain that permits the protein to be bound to an insoluble matrix. For example, GST-POSH fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with a potential labeled binding agent and incubated under conditions conducive to binding. Following incubation, the beads are washed to remove any unbound agent, and the matrix bead-bound label determined directly, or in the supernatant after the bound agent is dissociated.

In yet another embodiment, the POSH polypeptide and potential interacting polypeptide can be used to generate an interaction trap assay (see also, U.S. Patent NO: 5,283,317; Zervos et al. (1993) Cell 72:223-232; Madura et al. (1993) J Biol

Chem 268:12046-12054; Bartel et al. (1993) Biotechniques 14:920-924; and Iwabuchi et al. (1993) Oncogene 8:1693-1696), for subsequently detecting agents which disrupt binding of the proteins to one and other.

In particular, the method makes use of chimeric genes which express hybrid proteins. To illustrate, a first hybrid gene comprises the coding sequence for a DNA-binding domain of a transcriptional activator can be fused in frame to the coding sequence for a "bait" protein, e.g., a POSH polypeptide of sufficient length to bind to a potential interacting protein. The second hybrid protein encodes a transcriptional activation domain fused in frame to a gene encoding a "fish" protein, e.g., a potential interacting protein of sufficient length to interact with the POSH polypeptide portion of the bait fusion protein. If the bait and fish proteins are able to interact, e.g., form a POSH complex, they bring into close proximity the two domains of the transcriptional activator. This proximity causes transcription of a reporter gene which is operably linked to a transcriptional regulatory site responsive to the transcriptional activator, and expression of the reporter gene can be detected and used to score for the interaction of the bait and fish proteins.

One aspect of the present application provides reconstituted protein preparations including a POSH polypeptide and one or more interacting polypeptides.

In still further embodiments of the present assay, the POSH complex is generated in whole cells, taking advantage of cell culture techniques to support the subject assay. For example, as described below, the POSH complex can be constituted in a eukaryotic cell culture system, including mammalian and yeast cells. Often it will be desirable to express one or more viral proteins (e.g., Gag or Env) in such a cell along with a subject POSH polypeptide. It may also be desirable to infect the cell with a virus of interest. Advantages to generating the subject assay in an intact cell include the ability to detect inhibitors which are functional in an environment more closely approximating that which therapeutic use of the inhibitor would require, including the ability of the agent to gain entry into the cell. Furthermore, certain of the in vivo embodiments of the assay, such as examples given below, are amenable to high through-put analysis of candidate agents.

The components of the POSH complex can be endogenous to the cell selected to support the assay. Alternatively, some or all of the components can be derived from exogenous sources. For instance, fusion proteins can be introduced into the cell by recombinant techniques (such as through the use of an expression vector), as well as by microinjecting the fusion protein itself or mRNA encoding the fusion protein.

In many embodiments, a cell is manipulated after incubation with a candidate agent and assayed for a POSH or POSH-AP activity. In certain embodiments, a POSH-AP, such as PTPN12, is a tyrosine phosphatase. Tyrosine phosphatase activity may be assessed by incubating a cell lysate, which has optionally been treated with pervanadate to stimulate tyrosine phosphorylation, with a POSH-AP that has tyrosine phosphatase activity, immunoprecipitating the substrate protein and immunoblotting for the presence of phosphorylated tyrosine. Alternatively, tyrosine phosphatase activity may be assessed by the substrate trapping method. The substrate trapping method employs catalytically inactive mutants of a tyrosine phosphatase (e.g., a POSH-AP such as PTPN12). The catalytically inactive phosphatase mutant is immobilized on a solid matrix (e.g., AG25-protein A-Sepharose beads) and incubated with a substrate protein. The solid matrix to which the catalytically inactive phosphatase is bound is isolated and subjected to SDS-PAGE and immunoblotting for the presence of the substrate protein. The proteins employed in a phosphatase assay may optionally be purified proteins. (Lyons, PD et al (2001) J Biol Chem 246:24422-31; Garton, AJ et al (1996) Mol Cell Biol 16:6408-18).

In many embodiments, a cell is manipulated after incubation with a candidate agent and assayed for a POSH or POSH-AP activity. In certain embodiments a POSH or POSH-AP activity is represented by production of virus like particles. As demonstrated herein, an agent that disrupts POSH or POSH-AP activity can cause a decrease in the production of virus like particles. Other bioassays for POSH or POSH-AP activities may include apoptosis assays (e.g., cell survival assays, apoptosis reporter gene assays, etc.) and NF-kB nuclear localization assays (see e.g., Tapon et al. (1998) EMBO J. 17: 1395-1404). One apoptosis assay that may be used to assess TGN-associated protein activity is the TUNEL assay, which is used to

detect the presence of apoptotic cell death. In the TUNEL assay, the enzyme terminal deoxynucleotidyl transferase labels 3'-OH DNA ends (which are generated during apoptosis) with biotinylated nucleotides. The biotinylated nucleotides are then detected by immunoperoxidase staining. Another apoptosis assay that may be used to assess TGN-associated protein activity is the caspase assay, in which caspase activity is measured using a blue fluorescent substrate. Cleavage of the substrate by caspase 3 releases the fluorochrome, which then fluoresces green. An assay that may be employed to monitor cell proliferation associated with a TGN-associated protein is the MTT cell proliferation assay. The MTT cell proliferation assay is a colorimetric assay which measures the reduction of a tetrazolium component (MTT) into an insoluble formazan product by the mitochondria of viable cells. After incubation of the cells with the MTT reagent, a detergent solution is added to lyse the cells and solubilize the colored crystals. The samples may be read using an ELISA plate reader. The amount of color produced is directly proportional to the number of viable cells.

In certain embodiments, POSH or POSH-AP activities may include, without limitation, complex formation, ubiquitination and membrane fusion events (eg. release of viral buds or fusion of vesicles). POSH-AP activity may be assessed by the presence of phosphorylated substrate, such as, in the case of PKA, phosphorylated POSH. The interaction of POSH with a small GTPase such as Rac may also be indicative of the absence of phosphorylation of POSH by PKA. POSH complex formation may be assessed by immunoprecipitation and analysis of co-immunoprecipiated proteins or affinity purification and analysis of co-purified proteins. Fluorescence Resonance Energy Transfer (FRET)-based assays or other energy transfer assays may also be used to determine complex formation.

The effect of an agent that modulates the activity of POSH or a POSH-AP may be evaluated for effects on the trafficking of a protein through the secretory system. For example, the effects of the agent on the trafficking of the protein may be assessed by detecting the glycosylation of the protein in the presence and absence of the agent, for instance, through the use of antibodies specific for sugar moieties. For example, cell lysates from cells treated in the absence and presence of an agent that modulates the activity of POSH or a POSH-AP may be subjected to

immunoprecipitation and immunoblotting with antibodies directed to the glycoprotein of interest and the glycosylation state of the protein then compared.

Additional bioassays for assessing POSH and POSH-AP activities may include assays to detect the improper processing of a protein that is associated with a neurological disorder. One assay that may be used is an assay to detect the presence, including an increase or a decrease in the amount, of a protein associated with a neurological disorder. For example, the use of RNAi may be employed to knockdown the expression of a POSH or POSH-AP in cells (e.g., CHO cells or COS cells). The production of a secreted protein such as for example, amyloid beta, in the cell culture media, can then be assessed and compared to production of the secreted protein from control cells, which may be cells in which the POSH or POSH-AP activity has not been inhibited. The production of secreted proteins may be assessed, such as amyloid beta protein, which is associated with Alzheimer's disease. In some instances, a label may be incorporated into a secreted protein and the presence of the labeled secreted protein detected in the cell culture media. Proteins secreted from any cell type may be assessed, including for example, neural cells.

The effect of an agent that modulates the activity of POSH or a POSH-AP may be evaluated for effects on mouse models of various neurological disorders. For example, mouse models of Alzheimer's disease have been described. See, for example, United States Patent No. 5,612,486 for "Transgenic Animals Harboring APP Allele Having Swedish Mutation," Patent No. 5,850,003 (the '003 patent) for "Transgenic Rodents Harboring APP Allele Having Swedish Mutation," and United States Patent No. 5,455,169 entitled "Nucleic Acids for Diagnosing and Modeling Alzheimer's Disease". Mouse models of Alzheimer's disease tend to produce elevated levels of beta-amyloid protein in the brain, and the increase or decrease of such protein in response to treatment with a test agent may be detected. In some instances, it may also be desirable to assess the effects of a test agent on cognitive or behavioral characteristics of a mouse model for Alzheimer's disease, as well as mouse models for other neurological disorders.

In a further embodiment, transcript levels may be measured in cells having higher or lower levels of POSH or POSH-AP activity in order to identify genes that

are regulated by POSH or POSH-APs. Promoter regions for such genes (or larger portions of such genes) may be operatively linked to a reporter gene and used in a reporter gene-based assay to detect agents that enhance or diminish POSH- or POSH-AP-regulated gene expression. Transcript levels may be determined in any way known in the art, such as, for example, Northern blotting, RT-PCR, microarray, etc. Increased POSH activity may be achieved, for example, by introducing a strong POSH expression vector. Decreased POSH activity may be achieved, for example, by RNAi, antisense, ribozyme, gene knockout, etc.

In general, where the screening assay is a binding assay (whether protein-protein binding, agent-protein binding, etc.), one or more of the molecules may be joined to a label, where the label can directly or indirectly provide a detectable signal. Various labels include radioisotopes, fluorescers, chemiluminescers, enzymes, specific binding molecules, particles, e.g., magnetic particles, and the like. Specific binding molecules include pairs, such as biotin and streptavidin, digoxin and antidigoxin etc. For the specific binding members, the complementary member would normally be labeled with a molecule that provides for detection, in accordance with known procedures.

In further embodiments, the application provides methods for identifying targets for therapeutic intervention. A polypeptide that interacts with POSH or participates in a POSH-mediated process (such as viral maturation) may be used to identify candidate therapeutics. Such targets may be identified by identifying proteins that associated with POSH (POSH-APs) by, for example, immunoprecipitation with an anti-POSH antibody, in silico analysis of high-throughput binding data, two-hybrid screens, and other protein-protein interaction assays described herein or otherwise known in the art in view of this disclosure. Agents that bind to such targets or disrupt protein-protein interactions thereof, or inhibit a biochemical activity thereof may be used in such an assay. Targets that have been identified by such approaches include POSH-APs provided in Tables 7 and 8 and in Figure 36.

A variety of other reagents may be included in the screening assay. These include reagents like salts, neutral proteins, e.g., albumin, detergents, etc that are used to facilitate optimal protein-protein binding and/or reduce nonspecific or

background interactions. Reagents that improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti- microbial agents, etc. may be used. The mixture of components are added in any order that provides for the requisite binding. Incubations are performed at any suitable temperature, typically between 4 °C and 40 °C. Incubation periods are selected for optimum activity, but may also be optimized to facilitate rapid high-throughput screening.

In certain embodiments, a test agent may be assessed for antiviral or anticancer activity by assessing effects on an activity (function) of a POSH-AP. Activity (function) may be affected by an agent that acts at one or more of the transcriptional, translational or post-translational stages. For example, an siRNA directed to a POSH-AP encoding gene will decrease activity, as will a small molecule that interferes with a catalytic activity of a POSH-AP. In certain embodiments, the agent inhibits the activity of one or more polypeptides selected from among HERPUD1 and MSTP028.

7. Exemplary Nucleic Acids and Expression Vectors

In certain aspects, the application relates to nucleic acids encoding POSH polypeptides, such as, for example, SEQ ID Nos: 2, 5, 7, 9, 11, 26, 27, 28, 29 and 30. Nucleic acids of the application are further understood to include nucleic acids that comprise variants of SEQ ID Nos:1, 3, 4, 6, 8, 10, 31, 32, 33, 34, and 35. Variant nucleotide sequences include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as allelic variants; and will, therefore, include coding sequences that differ from the nucleotide sequence of the coding sequence designated in SEQ ID Nos:1, 3, 4, 6, 8 10, 31, 32, 33, 34, and 35, e.g., due to the degeneracy of the genetic code. In other embodiments, variants will also include sequences that will hybridize under highly stringent conditions to a nucleotide sequence of a coding sequence designated in any of SEQ ID Nos:1, 3, 4, 6, 8 10, 31, 32, 33, 34, and 35. Preferred nucleic acids of the application are human POSH sequences, including, for example, any of SEQ ID Nos: 1, 3, 4, 6, 31, 32, 33, 34, 35 and variants thereof and nucleic acids encoding an amino acid sequence selected from among SEQ ID Nos: 2, 5, 7, 26, 27, 28, 29 and 30.

In certain aspects, the application relates to nucleic acids encoding POSH-AP polypeptides. For example, POSH-APs of the disclosure are listed in Table 7. Nucleic acid sequences encoding these POSH-APs are provided in Figure 36. Additional examples of POSH-APs of the disclosure are provided in Table 8. In certain embodiments, variants will also include nucleic acid sequences that will hybridize under highly stringent conditions to a nucleotide sequence of a coding sequence of a POSH-AP. Preferred nucleic acids of the application are human POSH-AP sequences and variants thereof.

One of ordinary skill in the art will understand readily that appropriate stringency conditions which promote DNA hybridization can be varied. For example, one could perform the hybridization at 6.0 x sodium chloride/sodium citrate (SSC) at about 45 °C, followed by a wash of 2.0 x SSC at 50 °C. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 x SSC at 50 °C to a high stringency of about 0.2 x SSC at 50 °C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22 °C, to high stringency conditions at about 65 °C. Both temperature and salt may be varied, or temperature or salt concentration may be held constant while the other variable is changed. In one embodiment, the application provides nucleic acids which hybridize under low stringency conditions of 6 x SSC at room temperature followed by a wash at 2 x SSC at room temperature.

Isolated nucleic acids which differ from the POSH nucleic acid sequences or from the POSH-AP nucleic acid sequences due to degeneracy in the genetic code are also within the scope of the application. For example, a number of amino acids are designated by more than one triplet. Codons that specify the same amino acid, or synonyms (for example, CAU and CAC are synonyms for histidine) may result in "silent" mutations which do not affect the amino acid sequence of the protein. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences of the subject proteins will exist among mammalian cells. One skilled in the art will appreciate that these variations in one or more nucleotides (up to about 3-5% of the nucleotides) of the nucleic acids encoding a particular protein may exist among individuals of a given species due to natural

allelic variation. Any and all such nucleotide variations and resulting amino acid polymorphisms are within the scope of this application.

Optionally, a POSH or a POSH-AP nucleic acid of the application will genetically complement a partial or complete loss of function phenotype in a cell. For example, a POSH nucleic acid of the application may be expressed in a cell in which endogenous POSH has been reduced by RNAi, and the introduced POSH nucleic acid will mitigate a phenotype resulting from the RNAi. An exemplary POSH loss of function phenotype is a decrease in virus-like particle production in a cell transfected with a viral vector, optionally an HIV vector.

Another aspect of the application relates to POSH and POSH-AP nucleic acids that are used for antisense, RNAi or ribozymes. As used herein, nucleic acid therapy refers to administration or *in situ* generation of a nucleic acid or a derivative thereof which specifically hybridizes (e.g., binds) under cellular conditions with the cellular mRNA and/or genomic DNA encoding one of the POSH or POSH-AP polypeptides so as to inhibit production of that protein, e.g., by inhibiting transcription and/or translation. The binding may be by conventional base pair complementarity, or, for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix.

A nucleic acid therapy construct of the present application can be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of the cellular mRNA which encodes a POSH or POSH-AP polypeptide. Alternatively, the the construct is an oligonucleotide which is generated *ex vivo* and which, when introduced into the cell causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences encoding a POSH or POSH-AP polypeptide. Such oligonucleotide probes are optionally modified oligonucleotide which are resistant to endogenous nucleases, e.g., exonucleases and/or endonucleases, and is therefore stable *in vivo*. Exemplary nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patents 5,176,996; 5,264,564; and 5,256,775). Additionally, general approaches to constructing oligomers useful in nucleic acid therapy have been

reviewed, for example, by van der Krol et al., (1988) *Biotechniques* 6:958-976; and Stein et al., (1988) *Cancer Res* 48:2659-2668.

Accordingly, the modified oligomers of the application are useful in therapeutic, diagnostic, and research contexts. In therapeutic applications, the oligomers are utilized in a manner appropriate for nucleic acid therapy in general.

In another aspect of the application, the subject nucleic acid is provided in an expression vector comprising a nucleotide sequence encoding a POSH or POSH-AP polypeptide and operably linked to at least one regulatory sequence. Regulatory sequences are art-recognized and are selected to direct expression of the POSH or POSH-AP polypeptide. Accordingly, the term regulatory sequence includes promoters, enhancers and other expression control elements. Exemplary regulatory sequences are described in Goeddel; *Gene Expression Technology: Methods in Enzymology*, Academic Press, San Diego, CA (1990). For instance, any of a wide variety of expression control sequences that control the expression of a DNA sequence when operatively linked to it may be used in these vectors to express DNA sequences encoding a POSH or POSH-AP polypeptide. Such useful expression control sequences, include, for example, the early and late promoters of SV40, tet promoter, adenovirus or cytomegalovirus immediate early promoter, the lac system, the trp system, the TAC or TRC system, T7 promoter whose expression is directed by T7 RNA polymerase, the major operator and promoter regions of phage lambda , the control regions for fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α-mating factors, the polyhedron promoter of the baculovirus system and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof. It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transformed and/or the type of protein desired to be expressed. Moreover, the vector's copy number, the ability to control that copy number and the expression of any other protein encoded by the vector, such as antibiotic markers, should also be considered.

As will be apparent, the subject gene constructs can be used to cause expression of the POSH or POSH-AP polypeptides in cells propagated in culture,

e.g., to produce proteins or polypeptides, including fusion proteins or polypeptides, for purification.

This application also pertains to a host cell transfected with a recombinant gene including a coding sequence for one or more of the POSH or POSH-AP polypeptides. The host cell may be any prokaryotic or eukaryotic cell. For example, a polypeptide of the present application may be expressed in bacterial cells such as *E. coli*, insect cells (e.g., using a baculovirus expression system), yeast, or mammalian cells. Other suitable host cells are known to those skilled in the art. Accordingly, the present application further pertains to methods of producing the POSH or POSH-AP polypeptides. For example, a host cell transfected with an expression vector encoding a POSH polypeptide can be cultured under appropriate conditions to allow expression of the polypeptide to occur. The polypeptide may be secreted and isolated from a mixture of cells and medium containing the polypeptide. Alternatively, the polypeptide may be retained cytoplasmically and the cells harvested, lysed and the protein isolated. A cell culture includes host cells, media and other byproducts. Suitable media for cell culture are well known in the art. The polypeptide can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins, including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for particular epitopes of the polypeptide. In a preferred embodiment, the POSH or POSH-AP polypeptide is a fusion protein containing a domain which facilitates its purification, such as a POSH-GST fusion protein, POSH-intein fusion protein, POSH-cellulose binding domain fusion protein, POSH-polyhistidine fusion protein etc.

A recombinant POSH or POSH-AP nucleic acid can be produced by ligating the cloned gene, or a portion thereof, into a vector suitable for expression in either prokaryotic cells, eukaryotic cells, or both. Expression vehicles for production of a recombinant POSH or POSH-AP polypeptides include plasmids and other vectors. For instance, suitable vectors for the expression of a POSH polypeptide include plasmids of the types: pBR322-derived plasmids, pEMBL-derived plasmids, pEX-derived plasmids, pBTac-derived plasmids and pUC-derived plasmids for expression in prokaryotic cells, such as *E. coli.*

The preferred mammalian expression vectors contain both prokaryotic sequences to facilitate the propagation of the vector in bacteria, and one or more eukaryotic transcription units that are expressed in eukaryotic cells. The pcDNAI/amp, pcDNAI/neo, pRc/CMV, pSV2gpt, pSV2neo, pSV2-dhfr, pTk2, pRSVneo, pMSG, pSVT7, pko-neo and pHyg derived vectors are examples of mammalian expression vectors suitable for transfection of eukaryotic cells. Some of these vectors are modified with sequences from bacterial plasmids, such as pBR322, to facilitate replication and drug resistance selection in both prokaryotic and eukaryotic cells. Alternatively, derivatives of viruses such as the bovine papilloma virus (BPV-1), or Epstein-Barr virus (pHEBo, pREP-derived and p205) can be used for transient expression of proteins in eukaryotic cells. Examples of other viral (including retroviral) expression systems can be found below in the description of gene therapy delivery systems. The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells, as well as general recombinant procedures, see *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press, 1989) Chapters 16 and 17. In some instances, it may be desirable to express the recombinant POSH or POSH-AP polypeptide by the use of a baculovirus expression system. Examples of such baculovirus expression systems include pVL-derived vectors (such as pVL1392, pVL1393 and pVL941), pAcUW-derived vectors (such as pAcUW1), and pBlueBac-derived vectors (such as the ß-gal containing pBlueBac III).

Alternatively, the coding sequences for the polypeptide can be incorporated as a part of a fusion gene including a nucleotide sequence encoding a different polypeptide. This type of expression system can be useful under conditions where it is desirable, e.g., to produce an immunogenic fragment of a POSH or POSH-AP polypeptide. For example, the VP6 capsid protein of rotavirus can be used as an immunologic carrier protein for portions of polypeptide, either in the monomeric form or in the form of a viral particle. The nucleic acid sequences corresponding to the portion of the POSH or POSH-AP polypeptide to which antibodies are to be raised can be incorporated into a fusion gene construct which includes coding

sequences for a late vaccinia virus structural protein to produce a set of recombinant viruses expressing fusion proteins comprising a portion of the protein as part of the virion. The Hepatitis B surface antigen can also be utilized in this role as well. Similarly, chimeric constructs coding for fusion proteins containing a portion of a POSH polypeptide and the poliovirus capsid protein can be created to enhance immunogenicity (see, for example, EP Publication NO: 0259149; and Evans et al.,, (1989) *Nature* 339:385; Huang et al., (1988) *J. Virol.* 62:3855; and Schlienger et al., (1992) *J. Virol.* 66:2).

The Multiple Antigen Peptide system for peptide-based immunization can be utilized, wherein a desired portion of a POSH or POSH-AP polypeptide is obtained directly from organo-chemical synthesis of the peptide onto an oligomeric branching lysine core (see, for example, Posnett et al., (1988) *JBC* 263:1719 and Nardelli et al., (1992) *J. Immunol.* 148:914). Antigenic determinants of a POSH or POSH-AP polypeptide can also be expressed and presented by bacterial cells.

In another embodiment, a fusion gene coding for a purification leader sequence, such as a poly-(His)/enterokinase cleavage site sequence at the N-terminus of the desired portion of the recombinant protein, can allow purification of the expressed fusion protein by affinity chromatography using a Ni^{2+} metal resin. The purification leader sequence can then be subsequently removed by treatment with enterokinase to provide the purified POSH or POSH-AP polypeptide (e.g., see Hochuli et al., (1987) *J. Chromatography* 411:177; and Janknecht et al., *PNAS USA* 88:8972).

Techniques for making fusion genes are well known. Essentially, the joining of various DNA fragments coding for different polypeptide sequences is performed in accordance with conventional techniques, employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed to

generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al., John Wiley & Sons: 1992).

Table 2: Exemplary POSH nucleic acids

Sequence Name	Organism	Accession Number
cDNA FLJ11367 fis, clone HEMBA1000303	Homo sapiens	AK021429
Plenty of SH3 domains (POSH) mRNA	Mus musculus	NM_021506
Plenty of SH3s (POSH) mRNA	Mus musculus	AF030131
Plenty of SH3s (POSH) mRNA	Drosophila melanogaster	NM_079052
Plenty of SH3s (POSH) mRNA	Drosophila melanogaster	AF220364

Table 3: Exemplary POSH polypeptides

Sequence Name	Organism	Accession Number
SH3 domains-containing protein POSH	Mus musculus	T09071
plenty of SH3 domains	Mus musculus	NP_067481
Plenty of SH3s; POSH	Mus musculus	AAC40070
Plenty of SH3s	Drosophila melanogaster	AAF37265
LD45365p	Drosophila melanogaster	AAK93408
POSH gene product	Drosophila melanogaster	AAF57833

Plenty of SH3s	Drosophila melanogaster	NP_523776

In addition the following Tables provide the nucleic acid sequence and related SEQ ID NOs for domains of human POSH protein and a summary of POSH sequence identification numbers used in this application.

Table 4. Nucleic Acid Sequences and related SEQ ID NOs for domains in human POSH

Name of the sequence	Sequence	SEQ ID NO.
RING domain	TGTCCGGTGTGTCTAGAGCGCCTTGATGCTTCTGCGAAGGTCT TGCCTTGCCAGCATACGTTTTGCAAGCGATGTTTGCT GGGGATCGTAGGTTCTCGAAATGAACTCAGATGTCCCGAGT	31
1st SH_3 domain	CCATGTGCCAAAGCGTTATACAACTATGAAGGAAAAGAGCCTG GAGACCTTAAATTCAGCAAAGGCGACATCATCATTTT GCGAAGACAAGTGGATGAAAATTGGTACCATGGGGAAGTCAAT GGAATCCATGGCTTTTTCCCCACCAACTTTGTGCAGA TTATT	32
2nd SH_3 domain	CCTCAGTGCAAAGCACTTTATGACTTTGAAGTGAAAGACAAGG AAGCAGACAAAGATTGCCTTCCATTTGCAAAGGATGA TGTTCTGACTGTGATCCGAAGAGTGGATGAAAACTGGGCTGAA GGAATGCTGGCAGACAAAATAGGAATATTTCCAATTT CATATGTTGAGTTTAAC	33
3rd SH_3 domain	AGTGTGTATGTTGCTATATATCCATACACTCCTCGGAAAGAGG ATGAACTAGAGCTGAGAAAAGGGGAGATGTTTTTAGT GTTTGAGCGCTGCCAGGATGGCTGGTTCAAAGGGACATCCATG CATACCAGCAAGATAGGGGTTTTCCCTGGCAATTATG TGGCACCAGTC	34

4^{th} SH_3 domain	GAAAGGCACAGGGTGGTGGTTTCCTATCCTCCTCAGAGTGAGG CAGAACTTGAACTTAAAGAAGGAGATATTGTGTTTGT TCATAAAAAACGAGAGGATGGCTGGTTCAAAGGCACATTACAA CGTAATGGGAAAACTGGCCTTTTCCCAGGAAGCTTTG TGGAAAACA	35

Table 5. Summary of POSH sequence Identification Numbers

Sequence Information	Sequence Identification Number (SEQ ID NO)
Human POSH Coding Sequence	SEQ ID No: 1
Human POSH Amino Acid Sequence	SEQ ID No: 2
Human POSH cDNA Sequence	SEQ ID No: 3
5' cDNA Fragment of Human POSH	SEQ ID No: 4
N-terminus Protein Fragment of Human POSH	SEQ ID No: 5
3' mRNA Fragment of Human POSH	SEQ ID No: 6
C-terminus Protein Fragment of Human POSH	SEQ ID No: 7
Mouse POSH mRNA Sequence	SEQ ID No: 8
Mouse POSH Protein Sequence	SEQ ID No: 9
Drosophila melanogaster POSH mRNA Sequence	SEQ ID No: 10
Drosophila melanogaster POSH Protein Sequence	SEQ ID No: 11
Human POSH RING Domain Amino Acid Sequence	SEQ ID No: 26
Human POSH 1st SH_3 Domain Amino Acid Sequence	SEQ ID No: 27
Human POSH 2nd SH_3 Domain Amino Acid Sequence	SEQ ID No: 28
Human POSH 3rd SH_3 Domain Amino Acid Sequence	SEQ ID No: 29
Human POSH 4th SH_3 Domain Amino Acid Sequence	SEQ ID No: 30
Human POSH RING Domain Nucleic Acid Sequence	SEQ ID No: 31

Human POSH 1st SH_3 Domain Nucleic Acid Sequence	SEQ ID No: 32
Human POSH 2nd SH_3 Domain Nucleic Acid Sequence	SEQ ID No: 33
Human POSH 3rd SH_3 Domain Nucleic Acid Sequence	SEQ ID No: 34
Human POSH 4th SH_3 Domain Nucleic Acid Sequence	SEQ ID No: 35

8. Exemplary Polypeptides

In certain aspects, the present application relates to POSH polypeptides, which are isolated from, or otherwise substantially free of, other intracellular proteins which might normally be associated with the protein or a particular complex including the protein. In certain embodiments, POSH polypeptides have an amino acid sequence that is at least 60% identical to an amino acid sequence as set forth in any of SEQ ID Nos: 2, 5, 7, 9, 11, 26, 27, 28, 29 and 30. In other embodiments, the polypeptide has an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% or 100% identical to an amino acid sequence as set forth in any of SEQ ID Nos: 2, 5, 7, 9, 11, 26, 27, 28, 29 and 30.

In certain aspects, the application also relates to POSH-AP polypeptides (e.g., a POSH-AP provided in Table 7). Amino acid sequences of the POSH-APs listed in Table 7 are provided in Figure 36. Additional POSH-AP polypeptides are provided in Table 8. In certain embodiments, POSH-AP polypeptides have an amino acid sequence that is at least 60% identical to an amino acid sequence as set forth in Figure 36. In other embodiments, the POSH-AP polypeptide has an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% or 100% identical to an amino acid sequence as set forth in Figure 36.

Optionally, a POSH or POSH-AP polypeptide of the application will function in place of an endogenous POSH or POSH-AP polypeptide, for example by mitigating a partial or complete loss of function phenotype in a cell. For example, a POSH polypeptide of the application may be produced in a cell in which endogenous POSH has been reduced by RNAi, and the introduced POSH polypeptide will mitigate a phenotype resulting from the RNAi. An exemplary

POSH loss of function phenotype is a decrease in virus-like particle production in a cell transfected with a viral vector, optionally an HIV vector. In certain embodiments, a POSH polypeptide, when produced at an effective level in a cell, induces apoptosis.

In another aspect, the application provides polypeptides that are agonists or antagonists of a POSH or POSH-AP polypeptide. Variants and fragments of a POSH or POSH-AP polypeptide may have a hyperactive or constitutive activity, or, alternatively, act to prevent POSH or POSH-AP polypeptides from performing one or more functions. For example, a truncated form lacking one or more domain may have a dominant negative effect.

Another aspect of the application relates to polypeptides derived from a full-length POSH or POSH-AP polypeptide. Isolated peptidyl portions of the subject proteins can be obtained by screening polypeptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such polypeptides. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, any one of the subject proteins can be arbitrarily divided into fragments of desired length with no overlap of the fragments, or preferably divided into overlapping fragments of a desired length. The fragments can be produced (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments which can function as either agonists or antagonists of the formation of a specific protein complex, or more generally of a POSH:POSH-AP complex, such as by microinjection assays.

It is also possible to modify the structure of the POSH or POSH-AP polypeptides for such purposes as enhancing therapeutic or prophylactic efficacy, or stability (e.g., ex vivo shelf life and resistance to proteolytic degradation in vivo). Such modified polypeptides, when designed to retain at least one activity of the naturally-occurring form of the protein, are considered functional equivalents of the POSH or POSH-AP polypeptides described in more detail herein. Such modified polypeptides can be produced, for instance, by amino acid substitution, deletion, or addition.

For instance, it is reasonable to expect, for example, that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid (i.e,. conservative mutations) will not have a major effect on the biological activity of the resulting molecule. Conservative replacements are those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids are can be divided into four families (see, for example, Biochemistry, 2nd ed., Ed. by L. Stryer, W.H. Freeman and Co., 1981). Whether a change in the amino acid sequence of a polypeptide results in a functional homolog can be readily determined by assessing the ability of the variant polypeptide to produce a response in cells in a fashion similar to the wild-type protein. For instance, such variant forms of a POSH polypeptide can be assessed, e.g., for their ability to bind to another polypeptide, e.g., another POSH polypeptide or another protein involved in viral maturation. Polypeptides in which more than one replacement has taken place can readily be tested in the same manner.

This application further contemplates a method of generating sets of combinatorial mutants of the POSH or POSH-AP polypeptides, as well as truncation mutants, and is especially useful for identifying potential variant sequences (e.g., homologs) that are functional in binding to a POSH or POSH-AP polypeptide. The purpose of screening such combinatorial libraries is to generate, for example, POSH homologs which can act as either agonists or antagonist, or alternatively, which possess novel activities all together. Combinatorially-derived homologs can be generated which have a selective potency relative to a naturally occurring POSH or POSH-AP polypeptide. Such proteins, when expressed from recombinant DNA constructs, can be used in gene therapy protocols.

Likewise, mutagenesis can give rise to homologs which have intracellular half-lives dramatically different than the corresponding wild-type protein. For example, the altered protein can be rendered either more stable or less stable to proteolytic degradation or other cellular process which result in destruction of, or otherwise inactivation of the POSH or POSH-AP polypeptide of interest. Such homologs, and the genes which encode them, can be utilized to alter POSH or POSH-AP levels by modulating the half-life of the protein. For instance, a short

half-life can give rise to more transient biological effects and, when part of an inducible expression system, can allow tighter control of recombinant POSH or POSH-AP levels within the cell. As above, such proteins, and particularly their recombinant nucleic acid constructs, can be used in gene therapy protocols.

In similar fashion, POSH or POSH-AP homologs can be generated by the present combinatorial approach to act as antagonists, in that they are able to interfere with the ability of the corresponding wild-type protein to function.

In a representative embodiment of this method, the amino acid sequences for a population of POSH or POSH-AP homologs are aligned, preferably to promote the highest homology possible. Such a population of variants can include, for example, homologs from one or more species, or homologs from the same species but which differ due to mutation. Amino acids which appear at each position of the aligned sequences are selected to create a degenerate set of combinatorial sequences. In a preferred embodiment, the combinatorial library is produced by way of a degenerate library of genes encoding a library of polypeptides which each include at least a portion of potential POSH or POSH-AP sequences. For instance, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential POSH or POSH-AP nucleotide sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display).

There are many ways by which the library of potential homologs can be generated from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be carried out in an automatic DNA synthesizer, and the synthetic genes then be ligated into an appropriate gene for expression. The purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential POSH or POSH-AP sequences. The synthesis of degenerate oligonucleotides is well known in the art (see for example, Narang, SA (1983) Tetrahedron 39:3; Itakura et al., (1981) Recombinant DNA, Proc. 3rd Cleveland Sympos. Macromolecules, ed. AG Walton, Amsterdam: Elsevier pp273-289; Itakura et al., (1984) Annu. Rev. Biochem. 53:323; Itakura et al., (1984) Science 198:1056; Ike et al., (1983) Nucleic Acid Res. 11:477). Such techniques have been employed in the directed evolution of other proteins (see, for

example, Scott et al., (1990) Science 249:386-390; Roberts et al., (1992) PNAS USA 89:2429-2433; Devlin et al., (1990) Science 249: 404-406; Cwirla et al., (1990) PNAS USA 87: 6378-6382; as well as U.S. Patent Nos: 5,223,409, 5,198,346, and 5,096,815).

Alternatively, other forms of mutagenesis can be utilized to generate a combinatorial library. For example, POSH or POSH-AP homologs (both agonist and antagonist forms) can be generated and isolated from a library by screening using, for example, alanine scanning mutagenesis and the like (Ruf et al., (1994) Biochemistry 33:1565-1572; Wang et al., (1994) J. Biol. Chem. 269:3095-3099; Balint et al., (1993) Gene 137:109-118; Grodberg et al., (1993) Eur. J. Biochem. 218:597-601; Nagashima et al., (1993) J. Biol. Chem. 268:2888-2892; Lowman et al., (1991) Biochemistry 30:10832-10838; and Cunningham et al., (1989) Science 244:1081-1085), by linker scanning mutagenesis (Gustin et al., (1993) Virology 193:653-660; Brown et al., (1992) Mol. Cell Biol. 12:2644-2652; McKnight et al., (1982) Science 232:316); by saturation mutagenesis (Meyers et al., (1986) Science 232:613); by PCR mutagenesis (Leung et al., (1989) Method Cell Mol Biol 1:11-19); or by random mutagenesis, including chemical mutagenesis, etc. (Miller et al., (1992) A Short Course in Bacterial Genetics, CSHL Press, Cold Spring Harbor, NY; and Greener et al., (1994) Strategies in Mol Biol 7:32-34). Linker scanning mutagenesis, particularly in a combinatorial setting, is an attractive method for identifying truncated (bioactive) forms of POSH or POSH-AP polypeptides.

A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations and truncations, and, for that matter, for screening cDNA libraries for gene products having a certain property. Such techniques will be generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of POSH or POSH-AP homologs. The most widely used techniques for screening large gene libraries typically comprises cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the illustrative assays described below are amenable to high through-put analysis as

necessary to screen large numbers of degenerate sequences created by combinatorial mutagenesis techniques.

In an illustrative embodiment of a screening assay, candidate combinatorial gene products of one of the subject proteins are displayed on the surface of a cell or virus, and the ability of particular cells or viral particles to bind a POSH or POSH-AP polypeptide is detected in a "panning assay". For instance, a library of POSH variants can be cloned into the gene for a surface membrane protein of a bacterial cell (Ladner et al.,, WO 88/06630; Fuchs et al., (1991) Bio/Technology 9:1370-1371; and Goward et al., (1992) TIBS 18:136-140), and the resulting fusion protein detected by panning, e.g., using a fluorescently labeled molecule which binds the POSH polypeptide, to score for potentially functional homologs. Cells can be visually inspected and separated under a fluorescence microscope, or, where the morphology of the cell permits, separated by a fluorescence-activated cell sorter.

In similar fashion, the gene library can be expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences can be expressed on the surface of infectious phage, thereby conferring two significant benefits. First, since these phage can be applied to affinity matrices at very high concentrations, a large number of phage can be screened at one time. Second, since each infectious phage displays the combinatorial gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage can be amplified by another round of infection. The group of almost identical E. coli filamentous phages M13, fd, and fl are most often used in phage display libraries, as either of the phage gIII or gVIII coat proteins can be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle (Ladner et al., PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al., (1992) J. Biol. Chem. 267:16007-16010; Griffiths et al., (1993) EMBO J. 12:725-734; Clackson et al., (1991) Nature 352:624-628; and Barbas et al., (1992) PNAS USA 89:4457-4461).

The application also provides for reduction of the POSH or POSH-AP polypeptides to generate mimetics, e.g., peptide or non-peptide agents, which are able to mimic binding of the authentic protein to another cellular partner. Such mutagenic techniques as described above, as well as the thioredoxin system, are also

particularly useful for mapping the determinants of a POSH or POSH-AP polypeptide which participate in protein-protein interactions involved in, for example, binding of proteins involved in viral maturation to each other. To illustrate, the critical residues of a POSH or POSH-AP polypeptide which are involved in molecular recognition of a substrate protein can be determined and used to generate its derivative peptidomimetics which bind to the substrate protein, and by inhibiting POSH or POSH-AP binding, act to inhibit its biological activity. By employing, for example, scanning mutagenesis to map the amino acid residues of a POSH polypeptide which are involved in binding to another polypeptide, peptidomimetic compounds can be generated which mimic those residues involved in binding. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine (e.g., see Freidinger et al., in Peptides: Chemistry and Biology, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al., in Peptides: Chemistry and Biology, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gamma lactam rings (Garvey et al., in Peptides: Chemistry and Biology, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al., (1986) J. Med. Chem. 29:295; and Ewenson et al., in Peptides: Structure and Function (Proceedings of the 9th American Peptide Symposium) Pierce Chemical Co. Rockland, IL, 1985), b-turn dipeptide cores (Nagai et al., (1985) Tetrahedron Lett 26:647; and Sato et al., (1986) J Chem Soc Perkin Trans 1:1231), and b-aminoalcohols (Gordon et al., (1985) Biochem Biophys Res Commun 126:419; and Dann et al., (1986) Biochem Biophys Res Commun 134:71).

The following table provides the sequences of the RING domain and the various SH3 domains of POSH.

Table 6. Amino Acid Sequences and related SEQ ID NOs for domains in human POSH

Name of the sequence	Sequence	SEQ ID NO.
RING	CPVCLERLDASAKVLPCQHTFCKRCLLGIVGSRNELRCPEC	26

domain		
1st SH_3 domain	PCAKALYNYEGKEPGDLKFSKGDIIILRRQVDENWYHGEVNGIHGF FPTNFVQIIK	27
2nd SH_3 domain	PQCKALYDFEVKDKEADKDCLPFAKDDVLTVIRRVDENWAEGMLAD KIGIFPISYVEFNS	28
3rd SH_3 domain	SVYVAIYPYTPRKEDELELRKGEMFLVFERCQDGWFKGTSMHTSKI GVFPGNYVAPVT	29
4th SH_3 domain	ERHRVVVSYPPQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKT GLFPGSFVENI	30

The following table provides a list of selected POSH-APs and their related SEQ ID NOs.

Table 7 – Selected POSH APs

Protein	Protein Sequence (SEQ ID NO:)	mRNA Sequence (SEQ ID NO:)
ARF1	223	325-339
ARF5	224	340-344
ATP6V0C	225-226	345-351
CBL-B	361; 398; 227-230	353-360
CENTB1	231-232	37-47
DDEF1	233-237	48-54
EIF3S3	238	55-57
EPS8L2	239	58-60
GOCAP1	240-243	61-68
GOSR2	244-248	69-76
HERPUD1	249-252	77-86
HLA-A	253	87-88
HLA-B	254	89
MSTP028	255-256	90-94
PACS-1	362-366	95-100
PPP1CA	261-263; 395	101-110
PRKAR1A	264-265	111-122; 396-397
PTPN12	266-268	123-129
RALA	269-270	130-134
SIAH1	271-272	135-141
SMN1	273-275	142-146
SMN2	276-280	147-151
SNX1	281-286	152-161
SNX3	287-290	162-174

Protein	Protein Sequence (SEQ ID NO:)	mRNA Sequence (SEQ ID NO:)
SRA1	291-294	175-182
SYNE1	295-307	183-201
TTC3	308-312	202-207
UBE2N	313	208-210
UNC84B	314	211-213
VCY2IP1	315-323	214-222
SPG20	386-388	367-374
WASF1	389	375-376
HIP55	390-394	377-385

Table 8 below provides a list of POSH-APs that bound POSH in a 2-hybrid assay. Nucleic acid and amino acid sequences of the POSH-APs listed in Table 8 were filed in a U.S. provisional application filed in the name of Daniel N. Taglicht, Iris Alroy, Yuval Reiss, Liora Yaar, Danny Ben-Avraham, Shmuel Tuvia, and Tsvika Greener entitled "Posh Interacting Proteins and Related Methods", filed on March 2, 2004 (Attorney Docket No. PROL-P79-024), which Provisional Application is incorporated herein by reference in its entirety.

Table 8 – POSH-APs

Protein and Variant	Protein Sequence (public gi No.)	mRNA Sequence (public gi No.)
BCL9 – var 1	4757846	4757845
BRD4 – var 1	19718731	19718730
BRD4 – var 2	7657218	7657217
DRP2 – var 1	4503393	4503392
MAP1A – var 1	21536458	21536457
SH2D2A – var 1	4503633	31543620
BAT3 – var 1	18375630	18375633
BAT3 – var 2	18375634	18375631
BAT3 – var 3	*	18375629
BCAR1 – var 1	7656924	7656923
DAP – var 1	4758120	4758119
EVPL – var 1	4503613	4503612
FLJ13231 – var 1	38604073	38604072
FL53657 – var 1	13376230	13376229
HSPC142 – var 1	7661802	7661801
LOC118987 – var 1	29789403	31341089
NAP4 – var 1	2443367	2443366

Protein and Variant	Protein Sequence (public gi No.)	mRNA Sequence (public gi No.)
RBAF600 – var 1	24416002	24416001
XTP3TPB – var 1	20070264	20070263
Hs.31535 – var 1	37546355	37546354
ASF1B – var 1	8922549	8922548
ATP5A1 – var 1	4757810	23346425
C6 or fl 1– var 1	9954875	39725662
C6 or f60 – var 1	24431997	24431996
CDT1 – var 1	16418337	19923847
CIC – var 1	16507208	16507207
CLK2 – var 1	4557477	4557476
CLK2 – var 2	4502883	4502882
DNM2 – var 1	4826700	4826699
EEF1A1 – var 1	4503471	25453469
EIF4EBP1 – var 1	4758258	20070179
FLJ13479 – var 1	24432013	39725704
GC20 – var 1	5031711	5031710
GLUL – var 1	19923206	21361767
HEBP2 – var 1	7657603	7657602
ITGB- var 1	4504779	4504778
LAMA5 – var 1	21264602	21264601
LOC90987 – var 1	29734345	29734344
MRPL36 – var 1	23111040	20806105
Hs.380933 – var 1	30149441	37550602
NQO2 – var 1	4505417	4505416
PCBP1 – var 1	5453854	14141164
PCNT2 – var 1	22035674	35493922
PGD – var 1	984325	984324
RAP80 – var 1	21361593	21361592
RNH – var 1	21361547	21361546
RPL – var 1	4506597	15431291
RPS20 – var 1	4506697	14591915
RPS27A – var 1	4506713	27436941
SETDB1 – var 1	6912652	6912651
SF3A2 – var 1	21361376	32189413
UBB – var 1	11024714	22538474
ARHV – var 1	20070360	20070359
KIAA1111 – var 1	32698700	32698699
ZNF147 – var 1	4827065	15208652
PAWR – var 1	4505613	4505612
TPX2 – var 1	20127519	31542258
HSPA1B – var 1	4885431	26787974
DLG5 – var 1	3043690	3650451
DLG5 – var 2	28466997	28466996
DLG5 – var 3	3650452	16549841

Protein and Variant	Protein Sequence (public gi No.)	mRNA Sequence (public gi No.)
DLG5 – var 4	*	16807129
DLG5 – var 5	*	22539637
DLG5 – var 6	*	15929207
DLG5 – var 7	*	3043689
KIAA1598 – var 1	7023592	7023591
KIAA1598 – var 2	10047271	7018519
KIAA1598 – var 3	*	21314680
KIAA1598 – var 4	*	10047270
KIAA1598 – var 5	*	21755030
KIAA1598 – var 6	*	21755023
KIAA1598 – var 7	*	21754670
KIAA1598 – var 8	*	21750902
KIAA1598 – var 9	*	21749984
KIAA1598 – var 10	*	21749775
KIAA1598 – var 11	*	21749737
CGI-27 – var 1	7705720	23270696
CGI-27 – var 2	*	22902234
CGI-27 – var 3	*	17046302
CGI-27 – var 4	*	16553689
CGI-27 – var 5	*	10433504
CGI-27 – var 6	*	4680692
CGI-27 – var 7	*	20127543
BIA2 – var 1	5262640	5262639
BIA2 – var 2	21591225	21591224
BIA2 – var 3	*	21755615
COLIA1 – var 1	180392	407589
COLIA1 – var 2	180857	30015
COLIA1 – var 3	1418928	30092
COLIA1 – var 4	22328092	7209641
COLIA1 – var 5	762938	22328091
COLIA1 – var 6	30016	1418927
COLIA1 – var 7	407590	180856
COLIA1 – var 8	*	180391
COLIA1 – var 9	*	14719826
DKFZp761A052 – var 1	10434104	10434103
DKFZp761A052 – var 2	10439058	10439057
DKFZp761A052 – var 3	14602829	14602828
DKFZp761A052 – var 4	20380411	15079884
DKFZp761A052 – var 5	6808165	20380410
DKFZp761A052 – var 6	*	6808164
TLE1 – var 1	14603281	16041735
TLE1 – var 2	307510	14603280
TLE1 – var 3	*	307509
EGLN2 – var 1	8922130	23273571

Protein and Variant	Protein Sequence (public gi No.)	mRNA Sequence (public gi No.)
EGLN2 – var 2	12804603	10437903
EGLN2 – var 3	14547148	21733075
EGLN2 – var 4	18031805	21758140
EGLN2 – var 5	*	18677002
EGLN2 – var 6	*	18031804
EGLN2 – var 7	*	18141576
EGLN2 – var 8	*	14547147
EGLN2 – var 9	*	12804602
EGLN2 – var 10	*	10439822
EGLN2 – var 11	*	8922129
STC2 – var 1	3335144	3335143
STC2 – var 2	*	3702223
STC2 – var 3	*	4050037
STC2 – var 4	*	4104014
STC2 – var 5	*	13623494
STC2 – var 6	*	14042507
STC2 – var 7	*	14042032
STC2 – var 8	*	21755241
STC2 – var 9	*	21755207
STC2 – var 10	*	22761473
STC2 – var 11	*	12653744
OPTN – var 1	20149572	16550123
OPTN – var 2	21619683	3387890
OPTN – var 3	3329431	3127082
OPTN – var 4	3127083	3329430
OPTN – var 5	*	21619682
OPTN – var 6	*	18644681
OPTN – var 7	*	18644683
OPTN – var 8	*	18644685
OPTN – var 9	*	20149571
FLJ37147 – var 1	21753535	21753534
FLJ37147 – var 2	30153743	30153742
KHDRBS1 – var 1	21749696	189499
KHDRBS1 – var 2	1841747	12653852
KHDRBS1 – var 3	189500	17512262
KHDRBS1 – var 4	*	14714433
KHDRBS1 – var 5	*	1841746
KHDRBS1 – var 6	*	21749695
SLC2A1 – var 1	3387905	3387904
SLC2A1 – var 2	5730051	5730050
SLC2A1 – var 3	14268550	14268549
DKFZp434B1231 – var 1	6808117	6808116
NUMA1 – var 1	27694103	5453819
NUMA1 – var 2	35119	13278785

Protein and Variant	Protein Sequence (public gi No.)	mRNA Sequence (public gi No.)
NUMA1 – var 3	14249928	14249927
NUMA1 – var 4	13278786	15991876
NUMA1 – var 5	5453820	296118
NUMA1 – var 6	*	296119
NUMA1 – var 7	*	296120
NUMA1 – var 8	*	35118
NUMA1 – var 9	*	20073234
NUMA1 – var 10	*	22477305
NUMA1 – var 11	*	22749583
NUMA1 – var 12	*	27694102
HSPC016 – var 1	6841310	12654536
HSPC016 – var 2	12654537	6841309
HSPC016 – var 3	*	4679017
HSPC016 – var 4	*	10834763
UBC – var 1	5912028	3360475
UBC – var 2	340058	2647407
UBC – var 3	340068	24657521
UBC – var 4	14286308	21751700
UBC – var 5	15928840	21757163
UBC – var 6	16552475	21758959
UBC – var 7	*	16552474
UBC – var 8	*	15928839
UBC – var 9	*	14286307
UBC – var 10	*	12653358
UBC – var 11	*	10439801
UBC – var 12	*	340067
UBC – var 13	*	340057
UBC – var 14	*	5912027
ZFM1 – var 1	785999	785998
PIASY – var 1	14603164	3643110
PIASY – var 2	5533373	5533372
PIASY – var 3	24850133	10433892
PIASY – var 4	3643111	14603163
PIASY – var 5	*	20987516
PIASY – var 6	*	14709019
XM_208944 – var 1	30153743	30153742
J03930 – var 1	178442	178441
MT2A – var 1	187528	37120
MT2A – var 2	37121	263506
MT2A – var 3	*	13937856
MT2A – var 4	*	1495465
MT2A – var 5	*	187527
EWSR1 – var 1	7669490	21734132
EWSR1 – var 2	12653511	547565

Protein and Variant	Protein Sequence (public gi No.)	mRNA Sequence (public gi No.)
EWSR1 – var 3	15029675	21756356
EWSR1 – var 4	16552153	16551673
EWSR1 – var 5	16551674	16552152
EWSR1 – var 6	31280	15029674
EWSR1 – var 7	*	13435962
EWSR1 – var 8	*	12653510
EWSR1 – var 9	*	10439073
EWSR1 – var 10	*	7669489
MADH6 – var 1	2828712	1654326
MADH6 – var 2	2736316	20379504
MADH6 – var 3	1654327	2736315
MADH6 – var 4	*	2828711
MADH6 – var 5	*	15278059
THOC2 – var 1	20799318	10435649
THOC2 – var 2	10435650	20799317
THOC2 – var 3	*	7023224
ZNF151 – var 1	676873	2230870
ZNF151 – var 2	2230871	676872
DDX31 – var 1	10435700	14042193
DDX31 – var 2	10440004	15215272
DDX31 – var 3	20336298	16566549
DDX31 – var 4	16566550	20336297
DDX31 – var 5	15215273	20336296
DDX31 – var 6	14042194	10440003
DDX31 – var 7	*	10435699
POLR2J2 – var 1	11595478	21704271
POLR2J2 – var 2	21704274	21704270
POLR2J2 – var 3	19401711	19401710
POLR2J2 – var 4	14702175	21704273
POLR2J2 – var 5	21704272	16878085
POLR2J2 – var 6	*	11595475
POLR2J2 – var 7	*	11595477
POLR2J2 – var 8	*	11595473
BANF1 – var 1	3002951	11038645
BANF1 – var 2	4502389	13543576
BANF1 – var 3	*	14713907
BANF1 – var 4	*	3002950
BANF1 – var 5	*	4321975
BANF1 – var 6	*	3220254
CBX4 – var 1	1945453	1945452
CBX4 – var 2	15929016	2317722
CBX4 – var 3	2317723	15929015
ARIH2 – var 1	3925604	3925603
ARIH2 – var 2	9963793	3930777

Protein and Variant	Protein Sequence (public gi No.)	mRNA Sequence (public gi No.)
ARIH2 – var 3	12653307	3986675
ARIH2 – var 4	*	3986676
ARIH2 – var 5	*	3986677
ARIH2 – var 6	*	7328049
ARIH2 – var 7	*	6855602
ARIH2 – var 8	*	21749565
ARIH2 – var 9	*	33875424
ARIH2 – var 10	*	9963792
ARIH2 – var 11	*	5453556
ARIH2 – var 12	*	5817100
ARIH2 – var 13	*	3930775
SRPK2 – var 1	1857944	21752284
SRPK2 – var 2	23270876	21749007
SRPK2 – var 3	*	23270875
SRPK2 – var 4	*	1857943
SIAH2 – var 1	2673968	16549991
SIAH2 – var 2	2664283	34189635
SIAH2 – var 3	*	2664282
SIAH2 – var 4	*	2673967
KIAA0191 – var 1	27480017	29387261
KIAA0191 – var 2	1228035	10438300
KIAA0191 – var 3	29387262	1228034
KIAA0191 – var 4	*	21755057
KIAA0191 – var 5	*	27480016
KIAA0191 – var 6	*	19387907
KIAA0191 – var 7	*	15636651
KIAA0191 – var 8	*	23273514
PA1-RBP1 – var 1	5262551	22760761
PA1-RBP1 – var 2	4929579	20072477
PA1-RBP1 – var 3	12804377	17939456
PA1-RBP1 – var 4	12803339	18088243
PA1-RBP1 – var 5	14029171	16924316
PA1-RBP1 – var 6	18088244	33872286
PA1-RBP1 – var 7	22760762	14029170
PA1-RBP1 – var 8	*	33876749
PA1-RBP1 – var 9	*	12804376
PA1-RBP1 – var 10	*	4929578
PA1-RBP1 – var 11	*	4406639
PA1-RBP1 – var 12	*	5262550
FAT – var 1	2281025	1107686
FAT – var 2	1107687	15214611
FAT – var 3	*	2281024
FAT – var 4	*	598748
VCL – var 1	24657579	7669551

Protein and Variant	Protein Sequence (public gi No.)	mRNA Sequence (public gi No.)
VCL – var 2	340237	7669549
VCL – var 3	7669550	340236
VCL – var 4	*	21732673
VCL – var 5	*	15426616
VCL – var 6	*	246657578
SSR4 – var 1	15929882	30583222
SSR4 – var 2	13097213	1071680
SSR4 – var 3	*	22749791
SSR4 – var 4	*	21753447
SSR4 – var 5	*	16552704
SSR4 – var 6	*	15929881
SSR4 – var 7	*	13097212
SSR4 – var 8	*	2398656
PRDX5 – var 1	6166493	27484966
PRDX5 – var 2	6746355	9802047
PRDX5 – var 3	9802048	8745393
PRDX5 – var 4	27484967	6746354
PRDX5 – var 5	*	6563211
PRDX5 – var 6	*	6103723
PRDX5 – var 7	*	6166492
PRDX5 – var 8	*	6523288
PRDX5 – var 9	*	32455258
FLJ10120 – var 1	8922239	27469671
FLJ10120 – var 2	*	8922238
PROL4 – var 1	22208536	22208535
PROL4 – var 2	6005802	1050982
CL25084 – var 1	15341891	4406555
CL25084 – var 2	7023472	4406692
CL25084 – var 3	4406693	7023471
CL25084 – var 4	4406556	15341890
C11orf17 – var 1	22761313	21361869
C11orf17 – var 2	21105773	20149226
C11orf17 – var 3	20149225	20149224
C11orf17 – var 4	20149227	21105772
C11orf17 – var 5	21361870	21410957
C11orf17 – var 6	*	22761312
POLQ – var 1	3510695	13892060
POLQ – var 2	4163931	13892060
POLQ – var 3	13892061	4163930
POLQ – var 4	*	3510694
MBD2 – var 1	3170202	3800812
MBD2 – var 2	3800801	5817231
MBD2 – var 3	7710145	21595775
MBD2 – var 4	21595776	21464120

Protein and Variant	Protein Sequence (public gi No.)	mRNA Sequence (public gi No.)
MBD2 – var 5	*	21464121
MBD2 – var 6	*	3800800
MBD2 – var 7	*	3800792
MBD2 – var 8	*	3170201
FSTL1 – var 1	12658309	536897
FSTL1 – var 2	12652619	16924272
FSTL1 – var 3	*	33990756
FSTL1 – var 4	*	12658308
FSTL1 – var 5	*	10438502
FSTL1 – var 6	*	4884472

* denotes a polypeptide sequence that can be deduced from the corresponding mRNA sequence.

9. Effective Dose

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining The LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic induces are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the application, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (i.e., the concentration of the test compound which achieves a half-maximal

inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

10. Formulation and Use

Pharmaceutical compositions for use in accordance with the present application may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients. Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by, for example, injection, inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral or rectal administration.

An exemplary composition of the application comprises an RNAi mixed with a delivery system, such as a liposome system, and optionally including an acceptable excipient. In a preferred embodiment, the composition is formulated for topical administration for, e.g., herpes virus infections.

For such therapy, the compounds of the application can be formulated for a variety of loads of administration, including systemic and topical or localized administration. Techniques and formulations generally may be found in Remmington's Pharmaceutical Sciences, Meade Publishing Co., Easton, PA. For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, and subcutaneous. For injection, the compounds of the application can be formulated in liquid solutions, preferably in physiologically compatible buffers such as Hank's solution or Ringer's solution. In addition, the compounds may be formulated in solid form and redissolved or suspended immediately prior to use. Lyophilized forms are also included.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be

coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., ationd oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound. For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner. For administration by inhalation, the compounds for use according to the present application are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration bile salts and fusidic acid derivatives. in addition, detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories. For topical administration, the oligomers of the application are formulated into ointments, salves, gels, or creams as generally known in the art. A wash solution can be used locally to treat an injury or inflammation to accelerate healing.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

For therapies involving the administration of nucleic acids, the oligomers of the application can be formulated for a variety of modes of administration, including systemic and topical or localized administration. Techniques and formulations generally may be found in Remmington's Pharmaceutical Sciences, Meade Publishing Co., Easton, PA. For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, intranodal, and subcutaneous for injection, the oligomers of the application can be formulated in liquid solutions, preferably in physiologically compatible buffers such as Hank's solution or Ringer's solution. In addition, the oligomers may be formulated in solid form and

redissolved or suspended immediately prior to use. Lyophilized forms are also included.

Systemic administration can also be by transmucosal or transdermal means, or the compounds can be administered orally. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration bile salts and fusidic acid derivatives. In addition, detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories. For oral administration, the oligomers are formulated into conventional oral administration forms such as capsules, tablets, and tonics. For topical administration, the oligomers of the application are formulated into ointments, salves, gels, or creams as generally known in the art.

The application now being generally described, it will be more readily understood by reference to the following examples, which are included merely for purposes of illustration of certain aspects and embodiments of the present application, and are not intended to limit the application.

EXAMPLES

Example 1. Role of POSH in virus-like particle (VLP) budding

1. Objective:

Use RNAi to inhibit POSH gene expression and compare the efficiency of viral budding and GAG expression and processing in treated and untreated cells.

2. Study Plan:

HeLa SS-6 cells are transfected with mRNA-specific RNAi in order to knockdown the target proteins. Since maximal reduction of target protein by RNAi is achieved after 48 hours, cells are transfected twice – first to reduce target mRNAs, and subsequently to express the viral Gag protein. The second transfection is performed with pNLenv (plasmid that encodes HIV) and with low amounts of RNAi to maintain the knockdown of target protein during the time of gag expression and

budding of VLPs. Reduction in mRNA levels due to RNAi effect is verified by RT-PCR amplification of target mRNA.

3. Methods, Materials, Solutions

a. Methods

i. Transfections according to manufacturer's protocol and as described in procedure.

ii. Protein determined by Bradford assay.

iii. SDS-PAGE in Hoeffer miniVE electrophoresis system. Transfer in Bio-Rad mini-protean II wet transfer system. Blots visualized using Typhoon system, and ImageQuant software (ABbiotech)

b. Materials

Material	Manufacturer	Catalog #	Batch #
Lipofectamine 2000 (LF2000)	Life Technologies	11668-019	1112496
OptiMEM	Life Technologies	31985-047	3063119
RNAi Lamin A/C	Self	13	
RNAi TSG101 688	Self	65	
RNAi Posh 524	Self	81	
plenvl1 PTAP	Self	148	
plenvl1 ATAP	Self	149	
Anti-p24 polyclonal antibody	Seramun		A-0236/5-10-01
Anti-Rabbit Cy5 conjugated antibody	Jackson	144-175-115	48715
10% acrylamide Tris-Glycine SDS-PAGE gel	Life Technologies	NP0321	1081371
Nitrocellulose membrane	Schleicher & Schuell	401353	BA-83
NuPAGE 20X transfer buffer	Life Technologies	NP0006-1	224365
0.45μm filter	Schleicher &	10462100	CS1018-1

	Schuell		

c. Solutions

Lysis Buffer	Compound	Concentration
	Tris-HCl pH 7.6	50mM
	$MgCl_2$	15mM
	NaCl	150mM
	Glycerol	10%
	EDTA	1mM
	EGTA	1mM
	ASB-14 (add immediately before use)	1%
6X Sample Buffer	Tris-HCl, pH=6.8	1M
	Glycerol	30%
	SDS	10%
	DTT	9.3%
	Bromophenol Blue	0.012%
TBS-T	Tris pH=7.6	20mM
	NaCl	137mM
	Tween-20	0.1%

4. Procedure

a. Schedule

Day				
1	2	3	4	5
Plate cells	Transfection I (RNAi only)	Passage cells (1:3)	Transfection II (RNAi and pNlenv) (12:00, PM)	Extract RNA for RT-PCR (post transfection)

			Extract RNA for RT-PCR (pre-transfection)	Harvest VLPs and cells

b. Day 1

Plate HeLa SS-6 cells in 6-well plates (35mm wells) at concentration of 5 X10^5 cells/well.

c. Day 2

2 hours before transfection replace growth medium with 2 ml growth medium without antibiotics.

Transfection I:

Reaction	RNAi name	TAGDA#	Reactions	RNAi [nM]	RNAi [20μM] μl	A OPtiMEM (μl)	B LF2000 mix (μl)
1	Lamin A/C	13	2	50	12.5	500	500
2	Lamin A/C	13	1	50	6.25	250	250
3	TSG101 688	65	2	20	5	500	500
5	Posh 524	81	2	50	12.5	500	500

<u>Transfections:</u>

Prepare LF2000 mix: 250 μl OptiMEM + 5 μl LF2000 for each reaction. Mix by inversion, 5 times. Incubate 5 minutes at room temperature.

Prepare RNA dilution in OptiMEM (Table 1, column A). Add LF2000 mix dropwise to diluted RNA (Table 1, column B). Mix by gentle vortex. Incubate at room temperature 25 minutes, covered with aluminum foil.

Add 500 μl transfection mixture to cells dropwise and mix by rocking side to side.

Incubate overnight.

d. Day 3

Split 1:3 after 24 hours. (Plate 4 wells for each reaction, except reaction 2 which is plated into 3 wells.)

e. Day 4

2 hours pre-transfection replace medium with DMEM growth medium without antibiotics.

Transfection II

				A	B	C	D
RNAi name	TAG DA#	Plasmid	Reaction #	Plasmid for 2.4 µg (µl)	RNAi [20µM] for 10nM (µl)	OPtiMEM (µl)	LF2000 mix (µl)
Lamin A/C	13	PTAP	3	3.4	3.75	750	750
Lamin A/C	13	ATAP	3	2.5	3.75	750	750
TSG101 688	65	PTAP	3	3.4	3.75	750	750
Posh 524	81	PTAP	3	3.4	3.75	750	750

Prepare LF2000 mix: 250 µl OptiMEM + 5 µl LF2000 for each reaction. Mix by inversion, 5 times. Incubate 5 minutes at room temperature.

Prepare RNA+DNA diluted in OptiMEM (Transfection II, A+B+C)

Add LF2000 mix (Transfection II, D) to diluted RNA+DNA dropwise, mix by gentle vortex, and incubate 1h while protected from light with aluminum foil.

Add LF2000 and DNA+RNA to cells, 500µl/well, mix by gentle rocking and incubate overnight.

f. Day 5

Collect samples for VLP assay (approximately 24 hours post-transfection) by the following procedure (cells from one well from each sample is taken for RNA assay, by RT-PCR).

g. Cell Extracts

i. Pellet floating cells by centrifugation (5min, 3000 rpm at 4 °C), save supernatant (continue with supernatant immediately to step h), scrape remaining cells in the medium which remains in the well, add to the corresponding floating cell pellet and centrifuge for 5 minutes, 1800rpm at 4°C.

ii. Wash cell pellet twice with ice-cold PBS.

iii. Resuspend cell pellet in 100 µl lysis buffer and incubate 20 minutes on ice.

iv. Centrifuge at 14,000 rpm for 15 min. Transfer supernatant to a clean tube. This is the cell extract.

v. Prepare 10 µl of cell extract samples for SDS-PAGE by adding SDS-PAGE sample buffer to 1X, and boiling for 10 minutes. Remove an aliquot of the remaining sample for protein determination to verify total initial starting material. Save remaining cell extract at -80 °C.

h. Purification of VLPs from cell media

i. Filter the supernatant from step g through a 0.45m filter.

ii. Centrifuge supernatant at 14,000 rpm at 4 °C for at least 2 h.

iii. Aspirate supernatant carefully.

iv. Re-suspend VLP pellet in hot (100 °C warmed for 10 min at least) 1X sample buffer.

v. Boil samples for 10 minutes, 100 °C.

i. Western Blot analysis

i. Run all samples from stages A and B on Tris-Glycine SDS-PAGE 10% (120V for 1.5 h).

ii. Transfer samples to nitrocellulose membrane (65V for 1.5 h).

iii. Stain membrane with ponceau S solution.

iv. Block with 10% low fat milk in TBS-T for 1 h.

v. Incubate with anti p24 rabbit 1:500 in TBS-T o/n.

vi. Wash 3 times with TBS-T for 7 min each wash.

vii. Incubate with secondary antibody anti rabbit cy5 1:500 for 30 min.

viii. Wash five times for 10 min in TBS-T.

ix. View in Typhoon gel imaging system (Molecular Dynamics/APBiotech) for fluorescence signal.

Results are shown in Figures 11-13.

<u>Example 2. Exemplary POSH RT-PCR primers and siRNA duplexes</u>

RT-PCR primers

	Name	Position	Sequence
Sense primer	POSH=271	271	5' CTTGCCTTGCCAGCATAC 3' (SEQ ID NO:12)
Anti-sense primer	POSH=926c	926C	5' CTGCCAGCATTCCTTCAG 3' (SEQ ID NO:13)

siRNA duplexes:

siRNA No:	153	
siRNA Name:	POSH-230	
Position in mRNA	426-446	
Target sequence:	5' AACAGAGGCCTTGGAAACCTG 3'	SEQ ID NO: 14
siRNA sense strand:	5' dTdTCAGAGGCCUUGGAAACCUG 3'	SEQ ID NO: 15
siRNA anti-sense strand:	5'dTdTCAGGUUUCCAAGGCCUCUG 3'	SEQ ID NO: 16

siRNA No:	155	
siRNA Name:	POSH-442	
Position in mRNA	638-658	
Target sequence:	5' AAAGAGCCTGGAGACCTTAAA 3'	SEQ ID NO: 17
siRNA sense strand:	5' ddTdTAGAGCCUGGAGACCUUAAA 3'	SEQ ID NO: 18
siRNA anti-sense strand:	5' ddTdTUUUAAGGUCUCCAGGCUCU 3'	SEQ ID NO: 19

siRNA No:	157	
siRNA Name:	POSH-U111	
Position in mRNA	2973-2993	
Target sequence:	5' AAGGATTGGTATGTGACTCTG 3'	SEQ ID NO: 20
siRNA sense strand:	5' dTdTGGAUUGGUAUGUGACUCUG 3'	SEQ ID NO: 21
siRNA anti-sense strand:	5' dTdTCAGAGUCACAUACCAAUCC 3'	SEQ ID NO: 22

siRNA No:	159	
siRNA Name:	POSH-U410	
Position in mRNA	3272-3292	
Target sequence:	5' AAGCTGGATTATCTCCTGTTG 3'	SEQ ID NO: 23
siRNA sense strand:	5' ddTdTGCUGGAUUA UCUCCUGUUG 3'	SEQ ID NO: 24

siRNA anti-sense strand: 5' ddTdTCAACAGGAGAUAAUCCAGC 3' SEQ ID NO: 25

siRNA-No.: 187
siRNA Name: POSH-control
Position in mRNA: None. Reverse to #153
Target sequence: 5' AAGTCCAAAGGTTCCGGAGAC 3' SEQ ID NO: 36

3. Knock-down of hPOSH entraps HIV virus particles in intracellular vesicles.

HIV virus release was analyzed by electron microscopy following siRNA and full-length HIV plasmid (missing the envelope coding region) transfection. Mature viruses were secreted by cells transfected with HIV plasmid and non-relevant siRNA (control, lower panel). Knockdown of Tsg101 protein resulted in a budding defect, the viruses that were released had an immature phenotype (upper panel). Knockdown of hPOSH levels resulted in accumulation of viruses inside the cell in intracellular vesicles (middle panel). Results, shown in Figure 28, indicate that inhibiting hPOSH entraps HIV virus particles in intracellular vesicles. As accumulation of HIV virus particles in the cells accelerate cell death, inhibition of hPOSH therefore destroys HIV reservoir by killing cells infected with HIV.

Example 4. In-vitro assay of Human POSH self-ubiquitination

Recombinant hPOSH was incubated with ATP in the presence of E1, E2 and ubiquitin as indicated in each lane. Following incubation at 37 °C for 30 minutes, reactions were terminated by addition of SDS-PAGE sample buffer. The samples were subsequently resolved on a 10% polyacrylamide gel. The separated samples were then transferred to nitrocellulose and subjected to immunoblot analysis with an anti ubiquitin polyclonal antibody. The position of migration of molecular weight markers is indicated on the right.

Poly-Ub: Ub-hPOSHconjugates, detected as high molecular weight adducts only in reactions containing E1, E2 and ubiquitin. hPOSH-176 and hPOSH-178 are a short

and a longer derivatives (respectively) of bacterially expressed hPOSH; C, control E3.

Preliminary steps in a high-throughput screen

Materials

1. E1 recombinant from bacculovirus
2. E2 Ubch5c from bacteria
3. Ubiquitin
4. POSH #178 (1-361) GST fusion-purified but degraded
5. POSH # 176 (1-269) GST fusion-purified but degraded
6. hsHRD1 soluble ring containing region
5. Bufferx12 (Tris 7.6 40 mM, DTT 1mM, $MgCl_2$ 5mM, ATP 2uM)
6. Dilution buffer (Tris 7.6 40mM, DTT 1mM, ovalbumin 1ug/ul)

protocol

	0.1ug/ul	0.5ug/ul	5ug/ul	0.4ug/ul	2.5ug/u/	0.8ug/ul	
	E1	**E2**	**Ub**	**176**	**178**	**Hrd1**	**Bx12**
-E1 (E2+176)	-------	0.5	0.5	1	------	------	10
-E2 (E1+176)	1	------	0.5	1	------	------	9.5
-ub (E1+E2+176)	1	0.5	-------	1	------	------	9.5
E1+E2+176+Ub	1	0.5	0.5	1		------	9
-E1 (E2+178)	-------	0.5	0.5	------	1	-------	10
-E2 (E1+178)	1	-----	0.5	------	1	-------	9.5
-ub (E1+E2+178)	1	0.5	-----	------	1	-------	9.5
E1+E2+178+Ub	1	0.5	0.5	------	1	------1	9
Hrd1, E1+E2+Ub	1	0.5	0.5	------	-----	1	8.5

*

1. Incubate for 30 minutes at 37 °C.
2. Run 12% SDS PAGE gel and transfer to nitrocellulose membrane
3. Incubate with anti-Ubiquitin antibody.

Results, shown in Figure 19, demonstrate that human POSH has ubiquitin ligase activity.

Example 5. Co-immunoprecipitation of hPOSH with myc-tagged activated (V12) and dominant-negative (N17) Rac1

HeLa cells were transfected with combinations of myc-Rac1 V12 or N17 and hPOSHdelRING-V5. 24 hours after transfection (efficiency 80% as measured by GFP) cells were collected, washed with PBS, and swollen in hypotonic lysis buffer (10 mM HEPES pH=7.9, 15 mM KCl, 0.1 mM EDTA, 2 mM MgCl2, 1 mM DTT, and protease inhibitors). Cells were lysed by 10 strokes with dounce homogenizer and centrifuged 3000xg for 10 minutes to give supernatant (Fraction 1) and nucleii. Nucleii were washed with Fraction 2 buffer (0.2% NP-40, 10 mM HEPES pH=7.9, 40 mM KCl, 5% glycerol) to remove peripheral proteins. Nucleii were spun-down and supernatant collected (Fraction 2). Nuclear proteins were eluted in Fraction 3 buffer (20 mM HEPES pH=7.9, 0.42 M KCl, 25% glycerol, 0.1 mM EDTA, 2 mM $MgCl_2$, 1 mM DTT) by rotating 30 minutes in cold. Insoluble proteins were spun-down 14000xg and solubilized in Fraction 4 buffer (1% Fos-Choline 14, 50 mM HEPES pH=7.9, 150 mM NaCl, 10% glycerol, 1mM EDTA, 1.5 mM $MgCl_2$, 2 mM DTT). Half of the total extract was pre-cleared against Protein A sepharose for 1.5 hours and used for IP with 1 μg anti-myc (9E10, Roche 1-667-149) and Protein A sepharose for 2 hours. Immune complexes were washed extensively, and eluted in SDS-PAGE sample buffer. Gels were run, and proteins electro-transferred to nitrocellulose for immunoblot as in Figure 20. Endogenous POSH and transfected hPOSHdelRING-V5 are precipitated as a complex with Myc-Rac1 V12/N17. Results, shown in Figure 20, demonstrate that POSH co-immunoprecipitates with Rac1.

Example 6. POSH reduction results in decreased secretion of phospholipase D (PLD)

Hela SS6 cells (two wells of 6-well plate) were transfected with POSH siRNA or control siRNA (100 nM). 24 hours later each well was split into 5 wells of a 24-well plate. The next day cells were transfected again with 100 nM of either POSH siRNA or control siRNA. The next day cells were washed three times with 1xPBS and than 0.5 ml of PLD incubation buffer (118 mM NaCl, 6 mM KCl, 1 mM

$CaCl_2$, 1.2 mM MgSO4, 12.4 mM HEPES, pH7.5 and 1% fatty acid free bovine serum albumin) were added.

48 hours later medium was collected and centrifuged at 800xg for 15 minutes. The medium was diluted with 5xPLD reaction buffer (Amplex red PLD kit) and assayed for PLD by using the Amplex Red PLD kit (Molecular probes, A-12219). The assay results were quantified and presented below in as a bar graph. The cells were collected and lysed in 1% Triton X-100 lysis buffer (20 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EDTA, 1% Triton X-100 and 1x protease inhibitors) for 15 minutes on ice. Lysates were cleared by centrifugation and protein concentration was determined. There were equal protein concentrations between the two transfectants. Equal amount of extracts were immunoprecipitated with anti-POSH antibodies, separated by SDS-PAGE and immunoblotted with anti-POSH antibodies to assess the reduction of POSH levels. There was approximately 40% reduction in POSH levels (Figure 21).

Example 7. Effect of hPOSH on Gag-EGFP intracellular distribution

HeLa SS6 were transfected with Gag-EGFP, 24 hours after an initial transfection with either hPOSH–specific or scrambled siRNA (control) (100nM) or with plasmids encoding either wild type hPOSH or hPOSH C(12,55)A. Fixation and staining was preformed 5 hours after Gag-EGFP transfection. Cells were fixed, stained with Alexa fluor 647-conjugated Concanavalin A (ConA) (Molecular Probes), permeabilized and then stained with sheep anti-human TGN46. After the primary antibody incubation cells were incubated with Rhodamin-conjugated goat anti-sheep. Laser scanning confocal microscopy was performed on LSM510 confocal microscope (Zeiss) equipped with Axiovert 100M inverted microscope using x40 magnification and 1.3-numerical-aperture oil-immersion lens for imaging. For co-localization experiments, 10 optical horizontal sections with intervals of 1 μm were taken through each preparation (Z-stack). A single median section of each preparation is shown. See Figure 22.

Example 8. POSH-Regulated Intracellular Transport of Myristoylated Proteins

The localization of myristoylated proteins, Gag (see Figure 22), HIV-1 Nef, Src and Rapsyn, in cells depleted of hPOSH were analyzed by immunofluorescence. In control cells, HIV-1 Nef was found in a perinuclear region co-localized with hPOSH, indicative of a TGN localization (Figure 23). When hPOSH expression was reduced by siRNA treatment, Nef expression was weaker relative to control and nef lost its TGN, perinuclear localization. Instead it accumulated in punctated intracellular loci segregated from the TGN.

Src is expressed at the plasma membrane and in intracellular vesicles, which are found close to the plasma membrane (Figure 24, H187 cells). However, when hPOSH levels were reduced, Src was dispersed in the cytoplasm and loses its plasma membrane proximal localization detected in control (H187) cells (Figure 24, compare H153-1 and H187-2 panels).

Rapsyn, a peripheral membrane protein expressed in skeletal muscle, plays a critical role in organizing the structure of the nicotinic postsynaptic membrane (Sanes and Lichtman, Annu. Rev. Neurosci. 22: 389-442 (1999)). Newly synthesized Rapsyn associates with the TGN and than transported to the plasma membrane (Marchand et al., J. Neurosci. 22: 8891-01 (2002)). In hPOSH-depleted cells (H153-1) Rapsyn was dispersed in the cytoplasm, while in control cells it had a punctuated pattern and plasma membrane localization, indicating that hPOSH influences its intracellular transport (Figure 25).

Materials and Methods Used:

- Antibodies:

Src antibody was purchased from Oncogene research products(Darmstadt, Germany). Nef antibodies were pusrchased from ABI (Columbia, MA) and Fitzgerald Industries Interantional (Concord, MA). Alexa Fluor conjugated antibodies were pusrchased from Molecular Probes Inc. (Eugene, OR).

hPOSH antibody: Glutathione S-transferase (GST) fusion plasmids were constructed by PCR amplification of hPOSH codons 285-430. The amplified PCR products was cloned into pGEX-6P-2 (Amersham Pharmacia Biotech, Buckinghamshire, UK). The truncated hPOSH protein was generated in *E. coli*

BL21. Bacterial cultures were grown in LB media with carbenicillin (100 μg/ml) and recombinant protein production was induced with 1 mM IPTG for 4 hours at 30 °C. Cells were lysed by sonication and the recombinant protein was then isolated from the cleared bacterial lysate by affinity chromatography on a glutathione-sepharose resin (Amersham Pharmacia Biotech, Buckinghamshire, UK). The hPOSH portion of the fusion protein was then released by incubation with PreScission protease (Amersham Pharmacia Biotech, Buckinghamshire, UK) according to the manufacturer's instructions and the GST portion was then removed by a second glutathione-sepharose affinity chromatography. The purified partial hPOSH polypeptide was used to immunize New Zealand white rabbits to generate antibody 15B (Washington Biotechnology, Baltimore, Maryland).

- Construction of siRNA retroviral vectors:

hPOSH scrambled oligonucleotide (5'- CACACACTGCCG TCAACT GTTCAAGAGAC AGTTGACGGCAGTGTGTGTTTTTT -3'; and 5'- AATTAAAAAACACA CACTGCCGTCAACTGTC TCTTGAACAGTTGA CGGCAGTGTGTGGGCC -3') were annealed and cloned into the ApaI-EcoRI digested pSilencer 1.0-US (Ambion) to generate pSIL-scrambled. Subsequently, the U6-promoter and RNAi sequences were digested with BamHI, the ends filled in and the insert cloned into the Olil site in the retroviral vector, pMSVhyg (Clontech), generating pMSCVhyg-U6-scrambled. hPOSH oligonucleotide encoding RNAi against hPOSH (5'-AACAGAGGCCTTGGAAA CCTGGAAGC TTGCAGGTTT CCAAGGCCTCTGTT -3'; and 5'- GATCAACAGAG GCCTTGGAAACCTGC AAGCTTCCAGGTTTCCAA GGCCTCTGTT -3') were annealed and cloned into the BamHI-EcoRI site of pLIT-U6, generating pLIT-U6 hPOSH-230. pLIT-U6 is an shRNA vector containing the human U6 promoter (amplified by PCR from human genomic DNA with the primers, 5'-GGCCCACTAGTCA AGGTCG GGCA GGAAGA- 3' and 5'- GCCGAATT CAAAAAGGATC CGGCGATATCCGG TGTTTCGTCCTTTCCA -3') cloned into pLITMUS38 (New England Biolabs) digested with SpeI-EcoRI. Subsequently, the U6 promoter-hPOSH shRNA (pLIT-U6 hPOSH-230 digested with SnaBI and PvuI) was cloned into the Olil site of pMSVhyg (Clontech), generating pMSCVhyg U6-hPOSH-230.

• Generation of stable clones:

HEK 293T cells were transfected with retroviral RNAi plasmids (pMSCVhyg-U6-POSH-230 and pMSCVhyg-U6-scrambled and with plasmids encoding VSV-G and moloney gag-pol. Two days post transfection, medium containing retroviruses was collected and filtered and polybrene was added to a final concentration of 8μg/ml. This was used to infect HeLa SS6 cells grown in 60 mm dishes. Forty-eight hours post-infection cells were selected for RNAi expression by the addition of hygromycin to a final concentration of 300 μg/ml. Clones expressing RNAi against hPOSH were named H153, clones expressing scrambled RNAi were named H187.

• Transfection and immunofluorescent analysis:

Gag-EGFP experiments are described in Figure 22.

H153 or H187 cells were transfected with Src or Rapsyn-GFP (Image clone image: 3530551 or pNLenv-1). Eighteen hours post transfection cells were washed with PBS and incubated on ice with Alexa Fluor 647 conjugated Con A to label plasma membrane glycoproteins. Subsequently cells were fixed in 3% paraformaldehyde, blocked with PBS containing 4% bovine serum albumin and 1% gelatin. Staining with rabbit anti-Src, rabbit anti-hPOSH (15B) or mouse anti-nef was followed with secondary antibodies as indicated.

Laser scanning confocal microscopy was performed on LSM510 confocal microscope (Zeiss) equipped with Axiovert 100M inverted microscope using x40 magnification and 1.3-numerical-aperture oil-immersion lens for imaging. For co-localization experiments, 10 optical horizontal sections with intervals of 1 μm were taken through each preparation (Z-stack). A single median section of each preparation is shown.

Example 9. POSH Reduction by siRNA Abrogates West Nile Virus ("WNV") Infectivity.

HeLa SS6 cells were transfected with either control or POSH-specific siRNA. Cells were subsequently infected with WNV ($4x10^4$ PFU/well). Viruses

were harvested 24 hours and 48 hours post-infection, serially diluted, and used to infect Vero cells. As a control WNV ($4x10^4$ PFU/well), that was not passed through HeLa SS6 cells, was used to infect Vero cells. Virus titer was determined by plaque assay in Vero cells.

Virus titer was reduced by 2.5-log in cells treated with POSH-specific siRNA relative to cells transfected with control siRNA, thereby indicating that WNV requires POSH for virus secretion. See Figure 26.

Experimental Procedure:

- Cell culture, transfections and infection:

Hela SS6 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum and 100 units/ml penicillin and 100 μg/ml streptomycin. For transfections, HeLa SS6 cells were grown to 50% confluency in DMEM containing 10% FCS without antibiotics. Cells were then transfected with the relevant double-stranded siRNA (100 nM) using lipofectamin 2000 (Invitrogen, Paisley, UK). On the day following the initial transfection, cells were split 1:3 in complete medium and transfected with a second portion of double-stranded siRNA (50 nM). Six hours post-transfection medium was replaced and cells infected with WNV ($4x10^4$ PFU/well). Medium was collected from infected HeLa SS6 cells twenty-four and forty-eight post-infection (200 μl), serially diluted, and used to infect Vero cells. Virus titer was determined by plaque assay (Ben-Nathan D, Lachmi B, Lustig S, Feuerstien G (1991) Protection of dehydroepiandrosterone (DHEA) in mice ifected with viral encephalitis. Arch Viro; 120, 263-271).

<u>Example 10. Analysis of the effects of POSH knockdown on M-MuLV expression and budding</u>

Experimental Protocol:

Transfections:-

A day before transfection, Hela SS6 cells were plated in two 6 wells plates at $5 x10^5$ cells per well. 24 hours later the following transfections were performed: 4 wells were transfected with control siRNA and a plasmid encoding MMuLV.

4 wells were transfected with POSH siRNA and a plasmid encoding MMuLV.

1 well was a control without any siRNA or DNA transfected.

1 well was transfected with a plasmid encoding MMuLV.

For each well to be transfected 100 nM (12.5 µl) POSH siRNA or 100 nM (12.5 µl) control siRNA were diluted in 250 µl Opti-MEM (Invitrogen). Lipofectamin 2000 (5 µl) (Invitrogen, Cat. 11668-019) was mixed with 250 µl of OptiMEM per transfected well. The diluted siRNA was mixed with the lipofectamin 2000 mix and the solution incubated at room temperature for 30 min. The mixture was added directly to each well containing 2 ml DMEM +10% FBS (w/o antibiotics).

24 hours later, four wells of the same siRNA treatment were split to eight wells, and two wells without siRNA were split to four wells.

24 hours later all wells were transfected with 100 nM control siRNA or 100 nM POSH siRNA with or without a plasmid encoding MMuLV (see table below). 48 hours later virions and cells were harvested.

No of wells	RNAi	Amount of RNAi (µl) per well	Amount of DNA (µg) per well	The volume of DNA (µl) per well	Application
5	POSH 100 nM (1st and 2nd transfection)	12.5	MMuLV (2 µg)	10	4 wells for VLPs assay and 1 well for RT
5	Control 100 nM (1st and 2nd transfection)	12.5	MMuLV (2 µg)	10	4 wells for VLPs assay and 1 well for RT
1	-	-	-	10 µl H_2O	VLPs assay
1	-	-	MMuLV (2 µg)	10	VLPs assay

Steady state VLP assay

Cell extracts:-

1. Pellet floating cells by centrifugation (10 min, 500xg at 4 °C), save supernatant (continued at step 7), wash cells once, scrape cells in ice-cold 1xPBS, add to the corresponding cell pellet and centrifuge for 5 min 1800 rpm at 4 °C.
2. Wash cell pellet once with ice-cold 1xPBS.

3. Resuspend cell pellet in 150 µl 1% Triton X-100 lysis buffer (20 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EDTA, 1% Triton X-100 and 1x protease inhibitors) and incubate 20 minutes on ice.
4. Centrifuge at 14,000rpm for 15 min. Transfer supernatant to a clean tube.
5. Determine protein concentration by BCA.
6. Prepare samples for SDS-PAGE by adding 2 µl of 6xSB to 20 µg extract (add lysis buffer to a final volume of 12 µl), heat to 80 °C for 10 min.

Purification of virions from cell media

7. Filtrate the supernatant through a 0.45 µm filter.
8. Transfer 1500 µl of virions fraction to an ultracentrifuge tube (swinging rotor).
9. Add 300 µl of fresh sucrose cushion (20% sucrose in TNE) to the bottom of the tube.
10. Centrifuge supernatant at 35000 rpm at 4 °C for 2 hr.
11. Resuspend virion pellet in 50 µl hot 1x sample buffer each (samples 153-1, 2, 3, 187-1, 2, 3). Resuspend VLPs pellet (153-4, 5 and 187 4, 5) in 25 µl hot 1x sample buffer. Vortex shortly, transfer to an eppendorf tube, unite VLPs from wells 153-4+5 and 187- 4+5. Heat to 80 °C for 10 min.
12. Load equal amounts of VLPs relatively to cells extracts amounts.

Western Blot analysis

1. Separate all samples on 12% SDS-PAGE.
2. Transfer samples to nitrocellulose membrane (100V for 1.15 hr).
3. Dye membrane with ponceau solution.
4. Block with 10% low fat milk in TBS-T for 1 hour.
5. Incubate membranes with Goat anti p30 (81S-263) (1:5000) in 10% low fat milk in TBS-T over night at 4 °C. Incubate with secondary antibody rabbit anti goat-HRP 1:8000 for 60 min at room temperature.
6. Detect signal by ECL reaction.
7. Following the ECL detection incubate memebranes with Donkey anti rabbit Cy3 (Jackson Laboratories, Cat 711-165-152) 1:500 and detect signal by Typhoon scanning and quantitate.

Results:

As shown in Figure 27, POSH knockdown decreases the release of extracellular MMuLV particles.

Example 11. POSH Protein-protein interactions by yeast two hybrid assay

POSH-associated proteins were identified by using a yeast two-hybrid assay.

Procedure:

Bait plasmid (GAL4-BD) was transformed into yeast strain AH109 (Clontech) and transformants were selected on defined media lacking tryptophan. Yeast strain Y187 containing pre-transformed Hela cDNA prey (GAL4-AD) library (Clontech) was mated according to the Clontech protocol with bait containing yeast and plated on defined media lacking tryptophan, leucine, histidine and containing 2 mM 3 amino triazol. Colonies that grew on the selective media were tested for beta-galactosidase activity and positive clones were further characterized. Prey clones were identified by amplifying cDNA insert and sequencing using vector derived primers.

Bait:

Plasmid vector: pGBK-T7 (Clontech)

Plasmid name: pPL269- pGBK-T7 GAL4 POSHdR

Protein sequence: Corresponds to aa 53-888 of POSH (RING domain deleted)

RTLVGSGVEELPSNILLVRLLDGIKQRPWKPGPGGGSGTNCTNALRSQSSTVANCSSKDL
QSSQGGQQPRVQSWSPPVRGIPQLPCAKALYNYEGKEPGDLKFSKGDIIILRRQVDENWY
HGEVNGIHGFFPTNFVQIIKPLPQPPPQCKALYDFEVKDKEADKDCLPFAKDDVLTVIRR
VDENWAEGMLADKIGIFPISYVEFNSAAKQLIEWDKPPVPGVDAGECSSAAAQSSTAPKH
SDTKKNTKKRHSFTSLTMANKSSQASQNRHSMEISPPVLISSSNPTAAARISELSGLSCS
APSQVHISTTGLIVTPPPSSPVTTGPSFTFPSDVPYQAALGTLNPPLPPPPLLAATVLAS
TPPGATAAAAAAGMGPRPMAGSTDQIAHLRPQTRPSVYVAIYPYTPRKEDELELRKGEMF
LVFERCQDGWFKGTSMHTSKIGVFPGNYVAPVTRAVTNASQAKVPMSTAGQTSRGVTMVS
PSTAGGPAQKLQGNGVAGSPSVVPAAVVSAAHIQTSPQAKVLLHMTGQMTVNQARNAVRT
VAAHNQERPTAAVTPIQVQNAAGLSPASVGLSHHSLASPQPAPLMPGSATHTAAISISRA
SAPLACAAAAPLTSPSITSASLEAEPSGRIVTVLPGLPTSPDSASSACGNSSATKPDKDS
KKEKKGLLKLLSGASTKRKPRVSPPASPTLEVELGSAELPLQGAVGPELPPGGGHGRAGS
CPVDGDGPVTTAVAGAALAQDAFHRKASSLDSAVPIAPPPRQACSSLGPVLNESRPVVCE
RHRVVVSYPPQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKTGLFPGSFVENI

Library screened: Hela pretransformed library (Clontech).

POSH-APs identified by yeast two-hybrid assay are provided in Tables 7 and 8. Also, the nucleic acid and amino acid sequences of POSH-APs identified by yeast two-hybrid assay are provided in Figure 36. In addition, the nucleic acid and amino acid sequences of ARF1 and ARF5 are provided in Figure 36.

Example 12. Inhibition of PKA Kinase Activity Attenuates HIV-1 Virus Maturation

HeLa SS6 cells were transfected with pNLenv-1_{PTAP} or pNLenv-1_{ATAA} (L-domain mutant). Eighteen hours post-transfection, cells were transferred to 20 °C for two hours in order to inhibit transport of viral particles from the *trans*-Golgi (TGN) to the plasma membrane (PM). Subsequently, the PKA inhibitor, H89 (50 μM) (Biosource, Cat. No. PHZ1114) or DMSO were added to the cells and dishes were transferred to 37 °C to initiate transport from the TGN to the PM. Reverse transcriptase activity was assayed from virus-like-particles collected from cell supernatant twenty minutes later. H89 treatment resulted in complete inhibition of RT activity. Thus, demonstrating that PKA activity is required for HIV-1 viral maturation.

Materials and methods:

Cell culture and transfections

Hela SS6 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum and 100 units/ml penicillin and 100 μg/ml streptomycin. For transfections, HeLa SS6 cells were grown to 100% confluency in DMEM containing 10% FCS without antibiotics. Cells were then transfected with HIV-1_{NLenv1} (2 μg per 6-well) (Schubert et al., 1995).

Assays for virus release by RT activity

Virus and virus-like particle (VLP) release by reverse transcriptase activity was determined one day after transfection with the pro-viral DNA as previously described (Adachi et al., 1986; Fukumori et al., 2000; Lenardo et al., 2002). The culture medium of virus-expressing cells was collected and centrifuged at 500 x g

for 10 minutes. The resulting supernatant was passed through a 0.45 µm-pore filter and the filtrate was centrifuged at 14,000 x g for 2 hours at 4 °C. The resulting supernatant was removed and the viral-pellet was re-suspended in cell solubilization buffer (50 mM Tris-HCl, pH7.8, 80 mM potassium chloride, 0.75 mM EDTA and 0.5% Triton X-100, 2.5 mM DTT and protease inhibitors). The corresponding cells were washed three times with phosphate-buffered saline (PBS) and then solubilized by incubation on ice for 15 minutes in cell solubilization buffer. The cell detergent extract was then centrifuged for 15 minutes at 14,000 x g at 4 °C. The sample of the cleared extract (normally 1:10 of the initial sample) were resolved on a 12.5% SDS-polyacrylamide gel, then transferred onto nitrocellulose paper and subjected to immunoblot analysis with rabbit anti-CA antibodies. The CA was detected after incubation with a secondary anti-rabbit antibody conjugated to Cy5 (Jackson Laboratories, West Grove, Pennsylvania) and detected by fluorescence imaging (Typhoon instrument, Molecular Dynamics, Sunnyvale, California). The Pr55 and CA were then quantified by densitometry. A colorimetric reverse transcriptase assay (Roche Diagnostics GmbH, Mannenheim, Germany) was used to measure reverse transcriptase activity in VLP extracts. RT activity was normalized to amount of Pr55 and CA produced in the cells.

Example 13. hPOSH is phosphorylated by Protein kinase A (PKA)

PKA is a cAMP-dependent kinase. The holoenzyme is a tetramer of two catalytic subunits (cPKA) bound to two regulatory subunits PRKR1 or PRKR2. Activation proceeds by the cooperative binding of two cAMP molecules to each R subunit, which causes the dissociation of each active C subunit from the R subunit dimer. The consensus sequence for phosphorylation by the C subunit is, stringently, K/R-R-X-S/TY and less stringently, R-X-X-S/TY, where Y tends to be a hydrophobic residue. The intracellular localization of PKA is controlled thorough association with A-kinase-anchoring proteins (AKAPs). The regulatory subunit of protein kinase A (PRKR1A) was identified as a POSH interactor by yeast-two-hybrid screen, thereby implicating POSH as an AKAP.

Protein kinase A was demonstrated to be required for the budding of transport vesicles from the TGN (Muniz et al., 1997, Proc Natl Acad Sci U S A,

94:14461-6). Furthermore, it was demonstrated that an inhibitor of PKA, H89, is able to block HIV-1 release from cells (Cartier et al., 2003, J Biol Chem., 278:35211-9). Since POSH is localized at the TGN and is implicated as an AKAP, POSH may regulate PKA-mediated budding at the TGN of vesicles and HIV-1.

Applicants demonstrated that POSH is phosphorylated by PKA. Several putative PKA phosphorylation sites are found within hPOSH coding sequence (Figure 30). Phosphorylation of gravin, an AKAP, by PKA modulates its binding to the b2-adrenergic receptor. This serves to regulate the mobilization of gravin and PKA to the cell membrane and regulation of b2-AR activity by PKA. Two putative PKA sites are located in the putative-rac-binding region in POSH. Toward this end, POSH was subjected to in-vitro phosphorylation and binding to the small GTPase Rac1 (Figure 31). Indeed, only unphosphorylated POSH was able to bind activated, GTP-loaded, Rac1, demonstrating that phosphorylation regulates the binding of POSH to small GTPases, such as Rac1. GTPases of this sort family include TCL, TC10, Cdc42, Wrch-1, Rac2, Rac3 or RhoG (Aspenstrom et al., 2003, Biochem J., 377(Pt 2):327-37). Small GTPases of this sort are involved in protein trafficking in the secretory system, including the trafficking of viral proteins, such as those of HIV.

Materials and methods

PKA-dependent phosphorylation of hPOSH.

Bacterially expressed recombinant maltose-binding-protein (MBP)-hPOSH (3 μg) or GST-c-Cbl were incubated at 30oC for 30 minutes with (*) or without 10 ng PKA catalytic subunit (PKAc) in a buffer containing 40 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 4 mM ATP, 0.1 mg/ml BSA, 1 μM cAMP, 23 mM K_3PO_4, 7 nM DTT, and PKA peptide protection solution (Promega, Cat.No. V5340). The reaction was stopped by the addition of SDS-sample buffer, and boiling for 3 minutes. Samples were separated by SDS-PAGE on a 10% gel, and transferred to nitrocellulose and immunoblotted as detailed in the figure.

Binding of Rac1 to hPOSH

Bacterially expressed hPOSH (1 μg) or GST (1 μg) were phosphorylated as above. The reaction was terminated by the addition 0.5 ml of ice-cold 200 mM Tris-HCl pH 7.4, 5 mM EDTA. hPOSH and GST were then immobilized on NiNTA or reduced glutathione beads, respectively, by gentle mixing for 30 minutes. The immobilized proteins were washed three times with wash buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 5 mM MgCl2, 0.1 mM DTT). Recombinant Rac-1 (0.2 μg) (Sigma catalog # R3012) was incubated with or without 0.3 mM GTPγS (Sigma Cat. No. G8638) on ice for 15 minutes. The GTP/mock-loaded Rac-1 was then added to wash buffer (25 μl, final) and incubated for 30 minutes at 30 °C. The beads were then washed three times with wash buffer containing 0.1% Tween 20. Sample buffer was added to the bead pellet and boiled for 3 minutes. Immobilized and associating proteins were then separated by SDS-PAGE on a 12% gel and immunobloted with anti-Rac-1 (Santa Cruz Biotechnology, Cat. No. sc-217). Input is 0.25 μg of Rac-1.

Example 14. HERPUD1 Depletion by siRNA Reduces HIV Maturation.

Hela SS6 cells were transfeted with siRNA directed against HERPUD1 and with a plsmid encoding HIV proviral genome (pNLenv-1). Twenty four hours post-HIV transfection, virus-like particles (VLP) secreted into the medium were isolated and reverse transcriptase activity was determined. HIV release of active RT is an indication for a release of processed and mature virus. When the levels of HERPUD1 were reduced RT activity was inhibited by 80%, demonstrating the importance of HERPUD1 in HIV-maturation. See Figure 33.

Experimental Outline

- Cell culture and transfection:

HeLa SS6 were kindly provided by Dr. Thomas Tuschl (the laboratory of RNA Molecular Biology, Rockefeller University, New York, New York). Cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum and 100 U/ml penicillin and 100 μg/ml streptomycin. For transfections, HeLa SS6 cells were grown to 50% confluency in DMEM containing 10% FCS without antibiotics. Cells were then transfected with the relevant double-stranded siRNA (50-100nM) (HERPUD1: 5'-GGGAAGUUCUUCGGAACCUdTdT-3' and 5'-

dTdTCCCUUCAAGAAGCCUUGGA-5') using lipofectamin 2000 (Invitrogen, Paisley, UK). A day following the initial transfection cells were split 1:3 in complete medium and co-transfected 24 hours later with HIV-1NLenv1 (2 μg per 6-well) (Schubert et al., J. Virol. 72:2280-88 (1998)) and a second portion of double-stranded siRNA.

- Assay for virus release

Virus and virus-like particle (VLP) release was determined one day after transfection with the proviral DNA as previously described (Adachi et al., J. Virol. 59: 284-91 (1986); Fukumori et al., Vpr. Microbes Infect. 2: 1011-17 (2000); Lenardo et al., J. Virol. 76: 5082-93 (2002)). The culture medium of virus-expressing cells was collected and centrifuged at 500 x g for 10 minutes. The resulting supernatant was passed through a 0.45μm-pore filter and the filtrate was centrifuged at 14,000 x g for 2 hours at 4°C. The resulting supernatant was removed and the viral-pellet was re-suspended in SDS-PAGE sample buffer. The corresponding cells were washed three times with phosphate-buffered saline (PBS) and then solubilized by incubation on ice for 15 minutes in lysis buffer containing the following components: 50 mM HEPES-NaOH, (pH 7.5), 150 mM NaCl, 1.5 mM $MgCl_2$, 0.5% NP-40, 0.5% sodium deoxycholate, 1 mM EDTA, 1 mM EGTA and 1:200 dilution of protease inhibitor cocktail (Calbiochem, La Jolla, California). The cell detergent extract was then centrifuged for 15 minutes at 14,000 x g at 4°C. The VLP sample and a sample of the cleared extract (normally 1:10 of the initial sample) were resolved on a 12.5% SDS-polyacrylamide gel, then transferred onto nitrocellulose paper and subjected to immunoblot analysis with rabbit anti-CA antibodies. The CA was detected either after incubation with a secondary anti-rabbit horseradish peroxidase-conjugated antibody and detected by Enhanced Chemi-Luminescence (ECL) (Amersham Pharmacia) or after incubation with a secondary anti-rabbit antibody conjugated to Cy5 (Jackson Laboratories, West Grove, Pennsylvania) and detected by fluorescence imaging (Typhoon instrument, Molecular Dynamics, Sunnyvale, CA). The Pr55 and CA were then quantified by densitometry and the amount of released VLP was then determined by calculating the ratio between VLP-associated CA and intracellular CA and Pr55 as previously described (Schubert et al., J. Virol. 72:2280-88 (1998)).

- Analysis of reverse transcriptase activity in supernatants

RT activity was determined in pelleted VLP (see above) by using an RT assay kit (Roche, Germany; Cat.No. 1468120). Briefly, VLP pellets were resuspended in 40 µl RT assay lysis buffer and incubated at room temperature for 30 minutes. At the end of incubation 20 µl RT assay reaction mix was added to each sample and incubation continued at 37°C overnight. Samples (60 µl) were than transferred to MTP strip wells and incubated at 37°C for 1 hour. Wells were washed five times with wash buffer and DIG-POD added for a one-hour incubation at 37°C. At the end of incubation wells were washed five times with wash buffer and ABST substrate solution was added and incubated until color developed. The absorbance was read in an ELISA reader at 405 nm (reference wavelength 492 nm). The resulting signal intensity is directly proportional to RT activity; RT concentration was determined by plotting against a known amount of RT enzyme included in separate wells of the reaction.

Example 15. MSTP028 Reduction by siRNA Decreases HIV VLP Production.

This example demonstrates the effects of an siRNA-mediated decrease in MSTP028 expression on the production of HIV virus-like particles in HeLa cells. The effects were measured at steady state.

Experiments were performed according to two different protocols. Experiment 1 proceeded with a second transfection on day 3, while Experiment 2 involved an additional exchange of medium on day 3, and proceeded to the second transfection on day 4. The results from Experiment 1 are shown Figure 29A, and those for Experiment 2 are shown in Figure 29B.

Day 1: Preparing Cells

4.5X105 HeLa SS6 cells/well, were seeded in 1 x 6 well plates. Cells were seeded in transfection medium (growing medium free of antibiotics).

Materials:

Cat. No.	Manufacture	Reagent Name

D5796	Sigma	DMEM
04-121-1A	Beit Haemek	FCS
D8537	Sigma	PBS
P4333	Sigma	Pen/Strep
T4049	Sigma	0.25% Trypsin-EDTA

Day 2: Transfection

Materials:

Cat. No.	Manufacture	Reagent Name
11668-027	Invitrogen	LF2000 reagent
31985-047	GibcoBRL	OptiMEM

MSTP028 RNAi constructs:

	siRNA target sequence	Accession	Pos.
MST028	AAGTGCTCACCGACAGTGAAG	NM_031954	197
MST028	AAGATACTTATGAGCCTTTCT	NM_031954	392

Experimental and Control Conditions:

1- Control siRNA+ pNLEnv-1

2- POSH siRNA + pNLenv-1

3- MSTP028 siRNA + pNLenv-1

1. Two hours before transfection, replace cell media to 2ml/well complete DMEM without antibiotics.
2. siRNA dilution: for each transfection dilute 100 nm siRNA in 0.25 ml OptiMEM per well.
3. LF 2000 dilution: for each well dilute 5μl lipofectamine reagent in 0.25ml OptiMEM.
4. Incubate diluted siRNAs and LF 2000 for 5 minutes at RT.
5. Mix the diluted siRNAs with diluted LF2000 and incubated for 25 minutes at RT.

6. Add the mixture to the cells, 0.5 ml/well (drop wise) and incubate for 24 hours at 37°C in CO_2 incubator.

Transfections: for each well
(12.5 µl (siRNA)/ 0.25 ml OptiMEM) x 3
LF 2000 35 µl / 1.75 ml

Day 3:

Exp. 1: second transfection (as Day 4 below).
Exp. 2: Exchange medium.

Day 4:

Exp. 1: VLP assay (see below).
Exp. 2: Second transfection

1. Two hours before transfection, replace cell media to 2ml/well complete DMEM without antibiotics.
2. siRNA and DNA dilution: Prepare dilution of plasmid pNLenv-1 0.75 µg / well in 0.25 ml OptiMEM (total of 3 wells). Divide plasmid dilution to eppendorf tubes (0.25 ml each). To each tube add siRNA 40nM (2.5 µl).
3. LF 2000 dilution: for each well dilute 5µl lipofectamine reagent in 0.25ml OptiMEM.
4. Incubate diluted siRNAs and LF 2000 for 5 minutes at RT.
5. Mix the diluted siRNAs with diluted LF2000 and incubated for 1 hour at RT.
6. Add the mixture to the cells, 0.5 ml/well (drop wise) and incubate for 24 hours at 37°C in CO_2 incubator.

Day 5:

Exp. 2: VLP assay.

Solutions:

Lysis buffer

Tris-HCl pH 7.6	50mM
MgCl2	1.5mM
NaCl	150mM
Glycerol	10%
NP-40	0.5%
DOC	0.5%
EDTA	1mM
EGTA	1mM

Add PI_3C 1:200.

Steady state VLP assay

A. Cell extracts

1. Pellet floating cells by centrifugation (1min, 14000rpm at 40C), save supernatant (continue with supernatant immediately to step B), scrape cells in ice-cold PBS, add to the corresponding floated cell pellet and centrifuge for 5min 1800rpm at 40C.
2. Wash cell pellet once with ice-cold PBS.
3. Resuspend cell pellet (from 6 well) in 100 μl NP40-DOC lysis buffer and incubate 10 minutes on ice.
4. Centrifuge at 14,000rpm for 15min. Transfer supernatant to a clean eppendorf.
5. Prepare samples for SDS-PAGE by adding them sample buffer and boil for 10min - take the same volume for each reaction (15 μl).

B. Purification of VLP from cell media

1. Filtrate the supernatant through a 0.45μm filter.
2. Centrifuge supernatant at 14,000rpm at 40C for at least 2h.
3. Resuspend VLP pellet in 50 μl 1X sample buffer and boil for 10 min. Load 25 μl of each sample.

C. Western Blot analysis

1. Run all samples from stages A and B on Tris-Gly SDS-PAGE 12.5%.
2. Transfer samples to nitrocellulose membrane (100V for 1.15h.).
3. Dye membrane with ponceau solution.
4. Block with 10% low fat milk in TBS-t for 1h.
5. Incubate with anti p24 rabbit 1:500 in TBS-t 2 hour (room temperature) - overnight (40C).
6. Wash 3 times with TBS-t for 7min each wash.
7. Incubate with secondary antibody anti rabbit cy5 1:500 for 30min.
8. Wash five times for 10min in TBS-t
9. View in Typhoon for fluorescence signal (650).

Example 16. POSH-depleted cells have lower levels of Herp and it is not monoubiquitinated

POSH-depleted cells and their control counterparts were lysed and immunoblotted with anti-herp antibodies. Cells depleted of POSH (H153 RNAi stables cell lines) cells have lower levels of Herp compared with control cells (H187 RNAi) (Figure 34A panel A). When cells were trasnfected with a plasmid encoding flagged-tagged ubiquitin, and immunoprecipitated with anti-flag antibodies to immunoprecipitate ubiquitinated proteins, Herp was ubiquitinated only in H187 cells and not in H153 cells (Figure 34A panel B). When the aforementioned cells were transfected with Herp-encoding plasmid, exogenous herp levels were also reduced in H153 cells compared to H187 cells (Figure 34B panel A) and the ubiquitination of exogenous herp was reduced in the former cells, similar to endogenous Herp. The molecular weight of ubiquitinated Herp is as predicated to full-length Herp and does not seem as a high molecular weight smear, a characteristic of polyubiquitinated proteins. Thus POSH is responsible for the mono-ubiquitination of Herp, and in the absence of this modification herp is subjected to degradation, which may be mediated by the proteosome.

Materials and methods

Plasmid generation

Full-length Herp was cloned from image clone MGC:45131 IMAGE:5575914 (GeneBank Accesion BC032673) into pCMV-SPORT6.

Antibody production

Herp1 (amino acids 1 to 251) was amplified from a plasmid (3Gd4) obtained by yeast two hybrid screen for interactors of POSH. The amplified open reading frame was cloned into pGEX-6P, expressed in E. coli BL21 by induction with 1 mM IPTG and purified on glutathione-agarose. The purified protein was cleaved with Precision™ protease (Amersham Biosciences) and the GST moiety removed by glutathione chromatography. The protein was injected into rabbits (Washington Biotechnology) to produce anti-Herp1 sera.

Transfections and antibody detection

Twenty-four hours prior to transfection POSH-RNAi clones (H153) or control-RNAi clones (H187) cells were plated in 10 cm dishes in growth medium (DMEM containing 10% fetal calf serum without antibiotics). Cells were transfected with lipofectamin 2000 (Invitrogen Corporation) and either Herp-expression plasmid (2.5 μg) or empty vector (2.5 μg) and a vector encoding Flag-tagged ubiquitin (1 μg). Twenty-four hours post-trasnfection cells were lysed in lysis buffer (50 mM Tris-HCl, pH7.6, 1.5 mM MgCl2, 150 mM NaCL, 10% glycerol, 1 mM EDTA, 1 mM EGTA, 0.5% NP-40 and 0.5% sodium deoxycholate, containing protease inhibitors) and subjected to immunoprecipitation with anti-Flag antibodies (Sigma, F7425) to precipitate ubiquitinated proteins. Immunoprecipitated material and total cell lysates were separated on 10% SDS-PAGE and transferred to nitrocellulose membranes which were immunoblotted with anti-Herp antibodies.

Generation of H187 and H153 cell lines

To relieve the necessity for multiple transfections and to improve the reproducibility of hPOSH reduction, we have generated two cell lines, H187 and H153 constitutively expressing an integrated control and hPOSH siRNA (respectively).

Construction of shRNA retroviral vectors- hPOSH scrambled oligonucleotide (5'-CACACACTGCCGTCAACTGTTCAAGAGACAGTTGACGGCAGTGTGTGTTTTTT–3'; and 5'-AATTAAAAAACACACACTGCCGTCAACTGTCTCTTGAACAGTTGACGGCAGTGTGTGGGCC– 3') were annealed and cloned into the ApaI-EcoRI digested pSilencer 1.0-U6 (Ambion, Inc.) to generate pSIL-scrambled. Subsequently, the U6-promoter and RNAi sequences were digested with BamHI, and blunted by end filling. The insert was cloned into the OliI site in the retroviral vector, pMSCVhyg (BD Biosciences Clontech), generating pMSCVhyg-U6-scrambled. The hPOSH oligonucleotide encoding RNAi against hPOSH (5'-AACAGAGGCCTTGGAAACCTGGAAGCTTGCAGGTTTCCAAGGCCTCTGTT-3'; and 5'-GATCAACAGAGGCCTTGGAAACCTGCAAGCTTCCAGGTTTCCAAGGCCTCTGTT–3') were annealed and cloned into the BamHI-EcoRV site of pLIT-U6, generating pLIT-U6 hPOSH-230. The pLIT-U6 is an shRNA vector containing the human U6 promoter (amplified by PCR from human genomic DNA with the primers, 5'-GGCCCACTAGTCAAGGTCGGGCAGGAAGA-3' and 5'-GCCGAATTCAAAAAGGATCCGGCGATATCCGGTGTTTCGTCCTTTCCA-3') cloned into pLITMUS38 (New England Biolabs, Inc.) digested with SpeI-EcoRI. Subsequently, the U6 promoter-hPOSH shRNA (pLIT-U6 hPOSH-230 digested with SnaBI and PvuI) was cloned into the Olil site of pMSCVhyg (BD Biosciences Clontech) generating pMSCVhyg U6-hPOSH-230.

Recombinant retrovirus production- HEK 293T cells were transfected with retroviral RNAi plasmids (pMSCVhyg-U6-POSH-230 and pMSCVhyg-U6-scrambled and with plasmids encoding VSV-G and Moloney Gag-pol. Two days post-transfection, the retrovirus-containing medium was collected and filtered.

Infection and selection- Polybrene (Hexadimethrine bromide) (Sigma) (8μg/ml) was added to the filtered and the treated medium was subsequently used to infect HeLa SS6 cells. Forty-eight hours post-infection clones were selected for RNAi expression by the addition of hygromycin (300 μg/ml). Clones expressing the scrambled and the hPOSH RNAi were termed H187 and H153 (respectively).

Example 17. Inhibition of HBV production

HepG2.2.15 cells were plated on 9cm dishes and allowed to grow in 8% FCS for 5 days up to 70% confluence. After 5 days, cells were washed twice with PBS and re-supplied with fresh DMEM without FCS. In this medium, cells were treated every 24 hours with the depicted solutions (3μl solution/1ml medium) for another 4 days (4 treatments total). After 4 days, medium was collected from each plate, viruses were sedimented and analyzed.

As shown in Figure 35, lanes 7 and 8, compounds CAS number 14567-55-4 and CAS number 414908-38-0 inhibit HBV production at a concentration of 3μM. Detection of HBV proteins was performed essentially as described in Paran, N et al (2001) EMBO J 20(16):4443-4453.

INCORPORATION BY REFERENCE

All publications and patents mentioned herein are hereby incorporated by reference in their entirety as if each individual publication or patent was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control.

EQUIVALENTS

While specific embodiments of the subject applications have been discussed, the above specification is illustrative and not restrictive. Many variations of the applications will become apparent to those skilled in the art upon review of this specification and the claims below. The full scope of the applications should be determined by reference to the claims, along with their full scope of equivalents, and the specification, along with such variations.

PCT/US04/06308

UNITED STATES PATENT AND TRADEMARK OFFICE
DOCUMENT CLASSIFICATION BARCODE SHEET

New International Application

Claim(s)

7

Index 1.1.5.2
Version 1.0
Rev 12/06/01

Section

What Is Claimed:

1. An isolated, purified or recombinant complex comprising a POSH polypeptide and a POSH-associated protein (POSH-AP).

2. The complex of claim 1, wherein the POSH-AP comprises a polypeptide selected from the group consisting of: PKA, SNX1, SNX3, ATP6V0C, PTPN12, PPP1CA, GOSR2, CENTB1, DDEF1, ARF1, ARF5, PACS-1, EPS8L2, HERPUD1, UNC84B, MSTP028, GOCAP, EIF3S3, SRA1, CBL-B, RALA, SIAH1, SMN1, SMN2, SYNE1, TTC3, VCY2IP1 and UBE2N (UBC13).

3. The complex of claim 1, wherein the POSH-AP comprises a polypeptide selected from the group consisting of: ARHV (Chp), WASF1, HIP55, SPG20, HLA-A, and HLA-B.

4. The complex of any one of claims 1-3, wherein the POSH polypeptide is a human POSH polypeptide.

5. An isolated, purified or recombinant complex comprising HERPUD1 and a Ubiquitin ligase.

6. The complex of claim 5, wherein the Ubiquitin ligase is selected from the group consisting of: POSH, CBL-B, TTC3, and SIAH1.

7. A method for identifying an agent that modulates an activity of a POSH polypeptide or POSH-AP, the method comprising identifying an agent that disrupts a complex of any one of claims 1-3, wherein an agent that disrupts a complex of any of claims 1-3 is an agent that modulates an activity of the POSH polypeptide or the POSH-AP.

8. A method of identifying an antiviral agent, comprising:

(a) identifying a test agent that disrupts a complex comprising a POSH polypeptide and a POSH-AP; and

(b) evaluating the effect of the test agent on a function of a virus,

wherein an agent that inhibits a pro-infective or pro-replicative function of a virus is an antiviral agent.

9. The method of claim 8, wherein the POSH-AP is selected from the group consisting of: PKA, SNX1, SNX3, PTPN12, GOSR2, CENTB1, ARF1, ARF5, PACS-1, EPS8L2, HERPUD1, SMN1, SMN2, UNC84B, MSTP028, GOCAP, CBL-B, SYNE1, UBE2N (UBC13), SIAH1, TTC3, WASF1, HIP55, RALA, and SPG20.

10. The method of claim 8, wherein the virus is an envelope virus.

11. The method of claim 8, wherein the virus is a Human Immunodeficiency Virus.

12. The method of claim 8, wherein the virus is a West Nile Virus.

13. The method of claim 8, wherein the virus is Moloney Murine Leukemia Virus (MMuLV).

14. The method of claim 8, wherein evaluating the effect of the test agent on a function of the virus comprises evaluating the effect of the test agent on the budding or release of the virus or a virus-like particle.

15. A method of identifying an anti-apoptotic agent, comprising:

(a) identifying a test agent that disrupts a complex comprising a POSH polypeptide and a POSH-AP; and

(b) evaluating the effect of the test agent on apoptosis of a cell,

wherein an agent that decreases apoptosis of the cell is an anti-apoptotic agent.

16. A method of identifying an anti-cancer agent, comprising:

(a) identifying a test agent that disrupts a complex comprising a POSH polypeptide and a POSH-AP; and

(b) evaluating the effect of the test agent on proliferation or survival of a cancer cell,

wherein an agent that decreases proliferation or survival of a cancer cell is an anti-cancer cell.

17. The method of claim 16, wherein the POSH-AP is selected from the group consisting of: PKA, SNX1, PTPN12, PPP1CA, ARF1, ARF5, CENTB1, EPS8L2, EIF3S3, CBL-B, RALA, SIAH1, TTC3, ATP6V0C, and VCY2IP1.

18. The method of claim 16, wherein the cancer cell is a cell derived from a POSH-associated cancer.

19. A method of identifying an agent that inhibits trafficking of a protein through the secretory pathway, comprising:

(a) identifying a test agent that disrupts a complex comprising a POSH polypeptide and a POSH-AP; and

(b) evaluating the effect of the test agent on the trafficking of a protein through the secretory pathway

wherein an agent that disrupts localization of said POSH-AP is an agent that inhibits trafficking of a protein through the secretory pathway.

20. The method of claim 19, wherein step (b) comprises evaluating the effect of the test agent on the trafficking of a myristoylated protein through the secretory pathway.

21. The method of claim 19, wherein step (b) comprises evaluating the effect of the test agent on the trafficking of a viral protein through the secretory pathway.

22. The method of claim 19, wherein (b) comprises evaluating the effect of the test agent on the trafficking of a protein associated with a neurological disorder through the secretory pathway.

23. The method of claim 22, wherein the protein associated with a neurological disorder is amyloid beta precursor protein.

24. A method of identifying an agent that inhibits the progression of a neurological disorder, comprising:

(a) identifying a test agent that disrupts a complex comprising a POSH polypeptide and a POSH-AP; and

(b) evaluating the effect of the test agent on the trafficking of a protein through the secretory pathway

wherein an agent that disrupts localization of a POSH-AP is an agent that inhibits progression of a neurological disorder.

25. The method of claim 24, wherein the POSH-AP is selected from the group consisting of: HERPUD1, CBL-B, SIAH1, and TTC3.

26. The method of claim 25, wherein the POSH-AP is HERPUD1.

27. A method of identifying an agent that inhibits the progression of a neurological disorder, comprising:

(a) identifying a test agent that disrupts a complex comprising a POSH polypeptide and a POSH-AP; and

(b) evaluating the effect of the test agent on the ubiquitination of a protein.

28. The method of claim 27, wherein the POSH-AP is HERPUD1.

29. A method of treating a viral infection in a subject in need thereof, comprising administering an agent that inhibits a POSH-AP in an amount sufficient to inhibit the viral infection.

30. The method of claim 29, wherein the agent is selected from the group consisting of:

i) an agent that inhibits a kinase activity of the POSH-AP;

ii) an agent that inhibits expression of the POSH-AP;

iii) an agent that inhibits the ubiquitin ligase activity of the POSH-AP;

iv) an agent that inhibits the phosphatase activity of the POSH-AP;

v) an agent that inhibits the GTPase activity of the POSH-AP; and

vi) an agent that inhibits the ubiquitination of the POSH-AP.

31. The method of claim 29, wherein the POSH-AP comprises a polypeptide selected from the group consisting of: PKA, SNX1, SNX3, SMN1, SMN2, PTPN12, GOSR2, CENTB1, ARF1, ARF5, PACS-1, EPS8L2, HERPUD1, UNC84B, MSTP028, GOCAP, CBL-B, SYNE1, UBE2N (UBC13), SIAH1, TTC3, WASF1, HIP55, RALA, and SPG20.

32. The method of claim 31, wherein the POSH-AP comprises a polypeptide selected from the group consisting of: PKA, HERPUD1, MSTP028, CBL-B, and UBE2N (UBC13).

33. The method of claim 32, wherein said agent is selected from the group consisting of: an siRNA construct, a small molecule, an antibody, and an antisense construct.

34. The method of claim 33, wherein the agent is an siRNA construct comprising a nucleic acid sequence that hybridizes to an mRNA encoding the POSH-AP.

35. The method of claim 34, wherein the agent is an siRNA construct or an antisense construct that inhibits the expression of a polypeptide selected from the group consisting of PKA, HERPUD1, MSTP028, CBL-B, and UBE2N (UBC13).

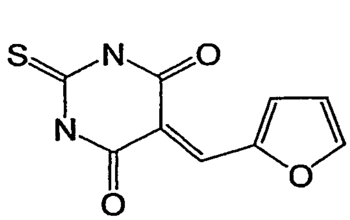
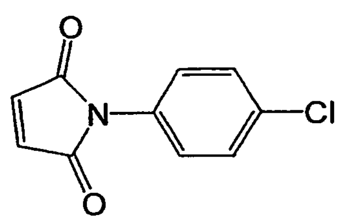
36. The method of claim 35, wherein the agent is an siRNA construct or an antisense construct that inhibits the expression of HERPUD1 or MSTP028.

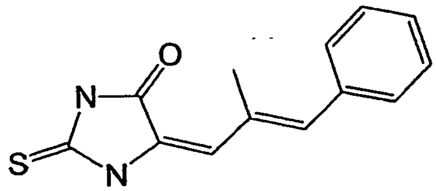
37. The method of claim 36, wherein the siRNA construct inhibits the expression of MSTP028.

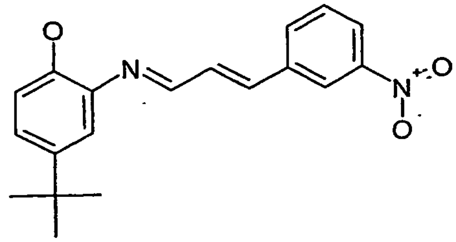
38. The method of claim 36, wherein the siRNA construct inhibits the expression of HERPUD1 and is selected from the group consisting of: 5'-GGGAAGUUCUUCGGAACCUdTdT-3' and 5'-dTdTCCCUUCAAGAAGCCUUGGA-5'.

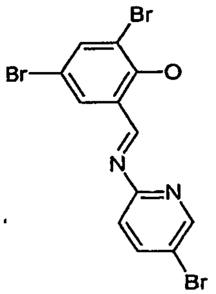
39. The method of claim 33, wherein the small molecule inhibitor is selected from among the following categories: adenosine cyclic monophosphorothioate, isoquinolinesulfonamide, piperazine, piceatannol, and ellagic acid.

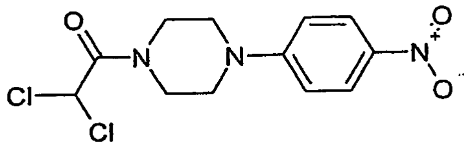
40. The method of claim 33, wherein the small molecule is selected from among:


;

;

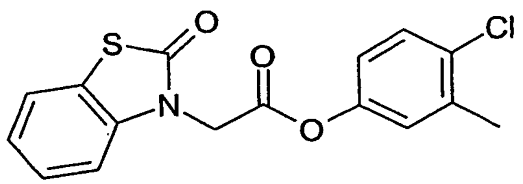
;

;

;

;

and

.

41. The method of claim 23, wherein the small molecule inhibits the ubiquitination of a POSH-AP.

42. The method claim 29, wherein the subject is infected with an envelope virus.

43. The method of claim 42, wherein the envelope virus is an HIV.

44. The method of claim 42, wherein the envelope virus is a WNV.

45. The method of claim 29, wherein the virus is a MMuLV.

46. Use of a protein kinase A inhibitor for the manufacture of a medicament for treatment of a viral infection.

47. Use of an inhibitor of HERPUD1 for the manufacture of a medicament for treatment of a viral infection.

48. Use of an inhibitor of MSTP028 for the manufacture of a medicament for treatment of a viral infection.

49. A packaged pharmaceutical for use in treating a viral infection, comprising:

(a) a pharmaceutical composition comprising an inhibitor of a POSH-AP and a pharmaceutically acceptable carrier; and

(b) instructions for use.

50. The packaged pharmaceutical of claim 49, wherein the viral infection is caused by an envelope virus.

51. A method for identifying an antiviral agent comprising:

(a) identifying a test agent that inhibits an activity of or expression of a POSH-AP; and

(b) evaluating an effect of the test agent on a function of a virus.

52. A method of evaluating an antiviral agent comprising:

(a) providing a test agent that inhibits an activity of or expression of a POSH-AP; and

(b) evaluating an effect of the test agent on a function of a virus.

53. The method of claim 51 or 52, wherein the virus is an envelope virus.

54. The method of claim 51 or 52, wherein the virus is a Human Immunodeficiency Virus.

55. The method of claim 51 or 52, wherein the virus is a West Nile Virus.

56. The method of claim 51 or 52, wherein the virus is a MMuLV.

57. The method of claim 51 or 52, wherein evaluating the effect of the test agent on a function of the virus comprises evaluating the effect of the test agent on the budding or release of the virus or a virus-like particle.

58. The method of claim 51 or 52, wherein the POSH-AP is selected from the group consisting of: PKA, SNX1, SNX3, PTPN12, GOSR2, SMN1, SMN2, CENTB1, ARF1, ARF5, PACS-1, EPS8L2, HERPUD1, UNC84B, MSTP028, GOCAP, CBL-B, SYNE1, UBE2N (UBC13), SIAH1, TTC3, WASF1, HIP55, RALA, and SPG20.

59. The method of claim 58, wherein the POSH-AP is HERPUD1.

60. The method of claim 58, wherein the POSH-AP is MSTP028.

61. The method of claim 51 or 52, wherein the test agent is selected from among: an antisense nucleic acid, an siRNA construct, a small molecule, an antibody and a polypeptide.

62. The method of claim 61, wherein the siRNA construct inhibits the expression of HERPUD1 and is selected from the group consisting of: 5'-GGGAAGUUCUUCGGAACCUdTdT-3' and 5'-dTdTCCCUUCAAGAAGCCUUGGA-5'.

63. A method of identifying an agent that modulates a POSH function, comprising:

a) identifying an agent that modulates a POSH-AP; and

b) testing the effect of the agent on a POSH function.

64. A method of evaluating an agent that modulates a POSH function, comprising:

a) providing an agent that modulates a POSH-AP; and

b) testing the effect of the agent on a POSH function.

65. The method of claim 64 or 65, wherein the POSH-AP comprises a polypeptide selected from the group consisting of: PKA, SNX1, SNX3, ATP6V0C, PTPN12, PPP1CA, GOSR2, CENTB1, DDEF1, ARF1, ARF5, PACS-1, EPS8L2, HERPUD1, UNC84B, MSTP028, GOCAP, EIF3S3, SRA1, CBL-B, RALA, SIAH1, SMN1, SMN2, SYNE1, TTC3, VCY2IP1 and UBE2N (UBC13).

66. The method of claim 64 or 65, wherein the POSH-AP comprises a polypeptide selected from the group consisting of: ARHV (Chp), WASF1, HIP55, SPG20, HLA-A, and HLA-B.

67. The method of claim 64 or 65, wherein testing the effect of the agent on a POSH function comprises testing the effect of the agent on the production of viral particles or virus like particles in a cell infected with an envelope virus.

68. The method of claim 64 or 65, wherein testing the effect of the agent on a POSH function comprises testing the effect of the agent on a POSH enzymatic activity.

69. The method of claim 68, wherein the POSH enzymatic activity is ubiquitin ligase activity.

70. The method of claim 64 or 65, wherein testing the effect of the agent on a POSH function comprises testing the effect of the agent on POSH-mediated localization or secretion of a protein.

71. The method of claim 64 or 65, wherein testing the effect of the agent on a POSH function comprises testing the effect of the agent on the interaction of POSH with a POSH-AP.

72. The method of claim 71, wherein the POSH-AP is a small GTPase.

73. The method of claim 72, wherein the small GTPase is selected from the group consisting of: ARF1, ARF5, and RALA.

74. The method of claim 64 or 65, wherein the test agent is selected from among: an antisense nucleic acid, an siRNA construct, a small molecule, an antibody and a polypeptide.

75. A method of identifying an agent that modulates a HERPUD1 function, comprising:

a) identifying an agent that modulates POSH; and

b) testing the effect of the agent on a HERPUD1 function.

76. A method of evaluating an agent that modulates an HERPUD1 function, comprising:

a) providing an agent that modulates POSH; and

b) testing the effect of the agent on a HERPUD1 function.

77. The method of claim 75 or 76, wherein testing the effect of the agent on a HERPUD1 function comprises contacting a cell with the agent and measuring the effect of the agent on ubiquitination of HERPUD1 in the cell.

78. A method of treating a viral infection in a subject in need thereof, comprising administering an agent that inhibits MSTP028 in an amount sufficient to inhibit viral infection.

79. The method of claim 78, wherein said agent is selected from the group consisting of: an siRNA construct, a small molecule, an antibody, and an antisense construct.

80. The method of claim 79, wherein the agent is an siRNA construct comprising a nucleic acid sequence that hybridizes to an mRNA encoding the MSTP028.

81. A method of inhibiting an activity of a POSH-AP in a cell, comprising contacting the cell with an inhibitor of POSH.

82. The method of claim 81, wherein the POSH-AP comprises a polypeptide selected from the group consisting of: PKA, SNX1, SNX3, ATP6V0C, PTPN12, PPP1CA, GOSR2, CENTB1, DDEF1, ARF1, ARF5, PACS-1, EPS8L2, HERPUD1, UNC84B, MSTP028, GOCAP, EIF3S3, SRA1, CBL-B, RALA, SIAH1, SMN1, SMN2, SYNE1, TTC3, VCY2IP1 and UBE2N (UBC13).

83. The method of claim 81, wherein the inhibitor of POSH is selected from among the following:

i) an agent that inhibits a POSH activity; and

ii) an agent that inhibits expression of a POSH.

84. The method of claim 83, wherein the POSH activity is ubiquitin ligase activity.

85. A method of treating a POSH-associated disease in a subject, comprising administering a POSH-AP inhibitor to a subject in need thereof.

86. The method of claim 85, wherein said POSH-AP inhibitor is an agent selected from the group consisting of:

i) an agent that inhibits a kinase activity of the POSH-AP;

ii) an agent that inhibits expression of the POSH-AP;

iii) an agent that inhibits the ubiquitin ligase activity of the POSH-AP;

iv) an agent that inhibits the phosphatase activity of the POSH-AP;

v) an agent that inhibits the GTPase activity of the POSH-AP; and

vi) an agent that inhibits the ubiquitination of the POSH-AP.

87. The method of claim 85, wherein the POSH-associated disease is a viral infection.

88. The method of claim 85, wherein the POSH-associated disease is a POSH-associated cancer.

89. The method of claim 85, wherein the POSH-associated disease is a POSH-associated neurological disorder.

90. A method of identifying an anti-viral agent, comprising:

a) forming a mixture comprising a POSH polypeptide, a POSH-AP and a test agent; and

b) detecting phosphorylation of the POSH polypeptide,

wherein an agent that inhibits phosphorylation of POSH is an anti-viral agent.

91. A method of identifying an anti-viral agent, comprising:

a) forming a mixture comprising a POSH polypeptide, a POSH-AP, ubiquitin and a test agent; and

b) detecting ubiquitination of the POSH-AP,

wherein an agent that inhibits ubiquitination of the POSH-AP is an anti-viral agent.

92. The method of claim 91, wherein the POSH-AP is HERPUD1.

93. A method of identifying a modulator of POSH, comprising:

a) forming a mixture comprising a POSH polypeptide, a POSH-AP and a test agent; and

b) detecting phosphorylation of the POSH polypeptide,

wherein an agent that alters phosphorylation of POSH is an agent that modulates POSH.

94. A method of identifying a modulator of POSH, comprising:

a) forming a mixture comprising a POSH polypeptide, a POSH-AP, ubiquitin and a test agent; and

b) detecting ubiquitination of the POSH-AP,

wherein an agent that inhibits ubiquitination of the POSH-AP is an agent that modulates POSH.

95. The method of claim 91, wherein the POSH-AP is HERPUD1.

96. A method of treating or preventing a POSH associated cancer in a subject comprising administering an agent that inhibits a POSH-AP to a subject in need thereof, wherein said agent treats or prevents cancer.

97. The method of claim 96, wherein the POSH-AP comprises a polypeptide selected from the group consisting of: PKA, SNX1, PTPN12, PPP1CA, CENTB1, ARF1, ARF5, EPS8L2, EIF3S3, CBL-B, RALA, SIAH1, TTC3, ATPV0C, and VCY2IP1.

98. The method of claim 96, wherein the cancer is associated with increased POSH expression.

99. A method of treating or preventing a POSH-associated neurological disorder in a subject comprising administering an agent that inhibits a POSH-AP to a subject in need thereof, wherein said agent treats or prevents the neurological disorder.

100. The method of claim 99, wherein the POSH-AP comprises a polypeptide selected from the group consisting of: PTPN12, DDEF1, EPS8L2, HERPUD1, GOCAP, CBL-B, SIAH1, SMN1, SMN2, TTC3, SPG20, SNX1, and ARF1.

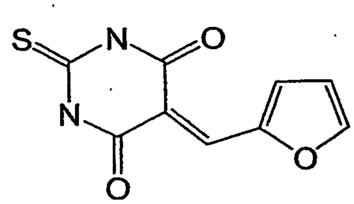
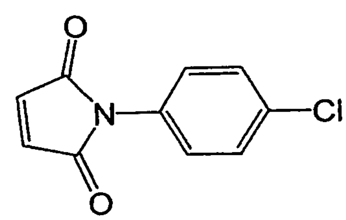
101. A method of treating a neurological disorder comprising administering an agent to a subject in need thereof, wherein said agent, inhibits the ubiquitin ligase activity of POSH.

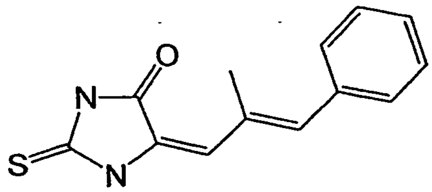
102. A method of treating a neurological disorder comprising administering an agent to a subject in need thereof, wherein said agent inhibits the ubquitination of a POSH-AP.

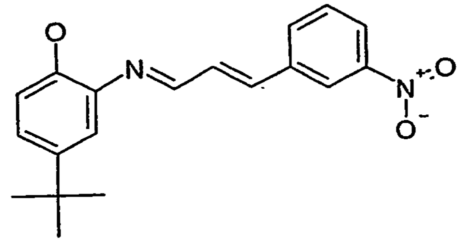
103. The method of claim 101 or claim 102, wherein the neurological disorder is selected from among: Alzheimer's disease, Parkinson's disease, Huntington's disease, schizophrenia, Niemann-Pick's disease, and prion-associated diseases.

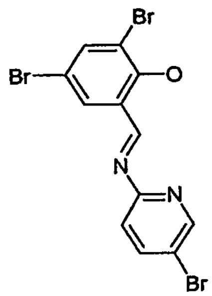
104. The use of an agent of claim 103, wherein the neurological disorder is Alzheimer's disease.

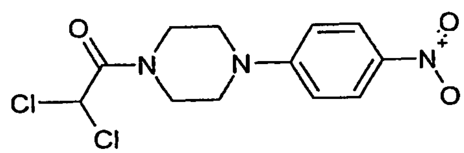
105. The method of claim 101 or claim 102, wherein said agent is selected from the group consisting of: an siRNA construct, a small molecule, an antibody, and an antisense construct.

106. The method of claim 105, wherein the small molecule is selected from among:  ;  ;

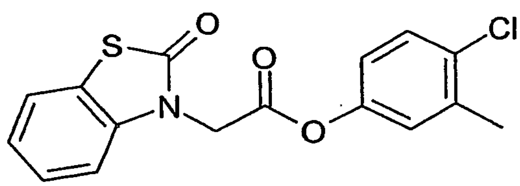
 ;

 ;

 ;

 ;

and



107. The method of claim 102, wherein the POSH-AP is HERPUD1.

108. The method of claim 61, wherein the siRNA construct inhibits the expression of MSTP028 and is selected from the group consisting of: 5'-AAGTGCTCACCGACAGTGAAG-3' and 5'-AAGATACTTATGAGCCTTTCT-3'.

POSH INTERACTING PROTEINS AND RELATED METHODS

ABSTRACT

The application provides novel complexes of POSH polypeptides and POSH-associated proteins. The application also provides methods and compositions for treating POSH-associated diseases such as viral disorders, cancer, and neurological disorders.

Figure 1: Human POSH Coding Sequence (SEQ ID NO:1) (part 1)

ATGGATGAATCAGCCTTGTTGGATCTTTTGGAGTGTCCGGTGTGTCTAGAGCGCCTTGATGCTTCTGCGA
AGGTCTTGCCTTGCCAGCATACGTTTTGCAAGCGATGTTTGCTGGGGATCGTAGGTTCTCGAAATGAACT
CAGATGTCCCGAGTGCAGGACTCTTGTTGGCTCGGGTGTCGAGGAGCTTCCCAGTAACATCTTGCTGGTC
AGACTTCTGGATGGCATCAAACAGAGGCCTTGGAAACCTGGTCCTGGTGGGGGAAGTGGGACCAACTGCA
CAAATGCATTAAGGTCTCAGAGCAGCACTGTGGCTAATTGTAGCTCAAAAGATCTGCAGAGCTCCCAGGG
CGGACAGCAGCCTCGGGTGCAATCCTGGAGCCCCCCAGTGAGGGGTATACCTCAGTTACCATGTGCCAAA
GCGTTATACAACTATGAAGGAAAAGAGCCTGGAGACCTTAAATTCAGCAAAGGCGACATCATCATTTTGC
GAAGACAAGTGGATGAAAATTGGTACCATGGGGAAGTCAATGGAATCCATGGCTTTTTCCCCACCAACTT
TGTGCAGATTATTAAACCGTTACCTCAGCCCCCACCTCAGTGCAAAGCACTTTATGACTTTGAAGTGAAA
GACAAGGAAGCAGACAAAGATTGCCTTCCATTTGCAAAGGATGATGTTCTGACTGTGATCCGAAGAGTGG
ATGAAAACTGGGCTGAAGGAATGCTGGCAGACAAAATAGGAATATTTCCAATTTCATATGTTGAGTTTAA
CTCGGCTGCTAAGCAGCTGATAGAATGGGATAAGCCTCCTGTGCCAGGAGTTGATGCTGGAGAATGTTCC
TCGGCAGCAGCCCAGAGCAGCACTGCCCCAAAGCACTCCGACACCAAGAAGAACACCAAAAAGCGGCACT
CCTTCACTTCCCTCACTATGGCCAACAAGTCCTCCCAGGCATCCCAGAACCGCCACTCCATGGAGATCAG
CCCCCCTGTCCTCATCAGCTCCAGCAACCCCACTGCTGCTGCACGGATCAGCGAGCTGTCTGGGCTCTCC
TGCAGTGCCCCTTCTCAGGTTCATATAAGTACCACCGGGTTAATTGTGACCCCGCCCCCAAGCAGCCCAG
TGACAACTGGCCCCTCGTTTACTTTCCCATCAGATGTTCCCTACCAAGCTGCCCTTGGAACTTTGAATCC
TCCTCTTCCACCACCCCCTCTCCTGGCTGCCACTGTCCTTGCCTCCACACCACCAGGCGCCACCGCCGCC
GCTGCTGCTGCTGGAATGGGACCGAGGCCCATGGCAGGATCCACTGACCAGATTGCACATTTACGGCCGC
AGACTCGCCCCAGTGTGTATGTTGCTATATATCCATACACTCCTCGGAAAGAGGATGAACTAGAGCTGAG
AAAAGGGGAGATGTTTTTAGTGTTTGAGCGCTGCCAGGATGGCTGGTTCAAAGGGACATCCATGCATACC
AGCAAGATAGGGGTTTTCCCTGGCAATTATGTGGCACCAGTCACAAGGGCGGTGACAAATGCTTCCCAAG
CTAAAGTCCCTATGTCTACAGCTGGCCAGACAAGTCGGGGAGTGACCATGGTCAGTCCTTCCACGGCAGG
AGGGCCTGCCCAGAAGCTCCAGGGAAATGGCGTGGCTGGGAGTCCCAGTGTTGTCCCCGCAGCTGTGGTA
TCAGCAGCTCACATCCAGACAAGTCCTCAGGCTAAGGTCTTGTTGCACATGACGGGGCAAATGACAGTCA
ACCAGGCCCGCAATGCTGTGAGGACAGTTGCAGCGCACAACCAGGAACGCCCCACGGCAGCAGTGACACC
CATCCAGGTACAGAATGCCGCCGGCCTCAGCCCTGCATCTGTGGGCCTGTCCCATCACTCGCTGGCCTCC
CCACAACCTGCGCCTCTGATGCCAGGCTCAGCCACGCACACTGCTGCCATCAGTATCAGTCGAGCCAGTG
CCCCTCTGGCCTGTGCAGCAGCTGCTCCACTGACTTCCCCAAGCATCACCAGTGCTTCTCTGGAGGCTGA
GCCCAGTGGCCGGATAGTGACCGTTCTCCCTGGACTCCCCACATCTCCTGACAGTGCTTCATCAGCTTGT
GGGAACAGTTCAGCAACCAAACCAGACAAGGATAGCAAAAAAGAAAAAAAGGGTTTGTTGAAGTTGCTTT
CTGGCGCCTCCACTAAACGGAAGCCCCGCGTGTCTCCTCCAGCATCGCCCACCCTAGAAGTGGAGCTGGG
CAGTGCAGAGCTTCCTCTCCAGGGAGCGGTGGGGCCCGAACTGCCACCAGGAGGTGGCCATGGCAGGGCA
GGCTCCTGCCCTGTGGACGGGGACGGACCGGTCACGACTGCAGTGGCAGGAGCAGCCCTGGCCCAGGATG
CTTTTCATAGGAAGGCAAGTTCCCTGGACTCCGCAGTTCCCATCGCTCCACCTCCTCGCCAGGCCTGTTC
CTCCCTGGGTCCTGTCTTGAATGAGTCTAGACCTGTCGTTTGTGAAAGGCACAGGGTGGTGGTTTCCTAT
CCTCCTCAGAGTGAGGCAGAACTTGAACTTAAAGAAGGAGATATTGTGTTTGTTCATAAAAAACGAGAGG
ATGGCTGGTTCAAAGGCACATTACAACGTAATGGGAAAACTGGCCTTTTCCCAGGAAGCTTTGTGGAAAA
CATATGA

Figure 2: Human POSH Amino Acid Sequence (SEQ ID NO:2) (part 2)

MDESALLDLLECPVCLERLDASAKVLPCQHTFCKRCLLGIVGSRNELRCPECRTLVGSGVEELPSNILLV
RLLDGIKQRPWKPGPGGGSGTNCTNALRSQSSTVANCSSKDLQSSQGGQQPRVQSWSPPVRGIPQLPCAK
ALYNYEGKEPGDLKFSKGDIIILRRQVDENWYHGEVNGIHGFFPTNFVQIIKPLPQPPPQCKALYDFEVK
DKEADKDCLPFAKDDVLTVIRRVDENWAEGMLADKIGIFPISYVEFNSAAKQLIEWDKPPVPGVDAGECS
SAAAQSSTAPKHSDTKKNTKKRHSFTSLTMANKSSQASQNRHSMEISPPVLISSSNPTAAARISELSGLS
CSAPSQVHISTTGLIVTPPPSSPVTTGPSFTFPSDVPYQAALGTLNPPLPPPPLLAATVLASTPPGATAA
AAAAGMGPRPMAGSTDQIAHLRPQTRPSVYVAIYPYTPRKEDELELRKGEMFLVFERCQDGWFKGTSMHT
SKIGVFPGNYVAPVTRAVTNASQAKVPMSTAGQTSRGVTMVSPSTAGGPAQKLQGNGVAGSPSVVPAAVV
SAAHIQTSPQAKVLLHMTGQMTVNQARNAVRTVAAHNQERPTAAVTPIQVQNAAGLSPASVGLSHHSLAS
PQPAPLMPGSATHTAAISISRASAPLACAAAAPLTSPSITSASLEAEPSGRIVTVLPGLPTSPDSASSAC
GNSSATKPDKDSKKEKKGLLKLLSGASTKRKPRVSPPASPTLEVELGSAELPLQGAVGPELPPGGGHGRA
GSCPVDGDGPVTTAVAGAALAQDAFHRKASSLDSAVPIAPPPRQACSSLGPVLNESRPVVCERHRVVVSY
PPQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKTGLFPGSFVENI

PCT/US04/06308

Figure 3: Human POSH cDNA Sequence (SEQ ID NO:3)

CTGAGAGACACTGCGAGCGGCGAGCGCGGTGGGGCCGCATCTGCATCAGCCGCCGCAGCCGCTGCGGGGC
CGCGAACAAAGAGGAGGAGCCGAGGCGCGAGAGCAAAGTCTGAAATGGATGTTACATGAGTCATTTTAAG
GGATGCACACAACTATGAACATTTCTGAAGATTTTTTCTCAGTAAAGTAGATAAAGATGGATGAATCAGC
CTTGTTGGATCTTTTGGAGTGTCCGGTGTGTCTAGAGCGCCTTGATGCTTCTGCGAAGGTCTTGCCTTGC
CAGCATACGTTTTGCAAGCGATGTTTGCTGGGGATCGTAGGTTCTCGAAATGAACTCAGATGTCCCGAGT
GCAGGACTCTTGTTGGCTCGGGTGTCGAGGAGCTTCCCAGTAACATCTTGCTGGTCAGACTTCTGGATGG
CATCAAACAGAGGCCTTGGAAACCTGGTCCTGGTGGGGAAGTGGACCAACTGCACAAATGCATTAAGG
TCTCAGAGCAGCACTGTGGCTAATTGTAGCTCAAAAGATCTGCAGAGCTCCCAGGGCGGACAGCAGCCTC
GGGTGCAATCCTGGAGCCCCCCAGTGAGGGGTATACCTCAGTTACCATGTGCCAAAGCGTTATACAACTA
TGAAGGAAAAGAGCCTGGAGACCTTAAATTCAGCAAAGGCGACATCATCATTTTGCGAAGACAAGTGGAT
GAAAATTGGTACCATGGGGAAGTCAATGGAATCCATGGCTTTTTCCCCACCAACTTTGTGCAGATTATTA
AACCGTTACCTCAGCCCCCACCTCAGTGCAAAGCACTTTATGACTTTGAAGTGAAAGACAAGGAAGCAGA
CAAAGATTGCCTTCCATTTGCAAAGGATGATGTTCTGACTGTGATCCGAAGAGTGGATGAAAACTGGGCT
GAAGGAATGCTGGCAGACAAAATAGGAATATTTCCAATTTCATATGTTGAGTTTAACTCGGCTGCTAAGC
AGCTGATAGAATGGGATAAGCCTCCTGTGCCAGGAGTTGATGCTGGAGAATGTTCCTCGGCAGCAGCCCA
GAGCAGCACTGCCCCAAAGCACTCCGACACCAAGAAGAACACCAAAAAGCGGCACTCCTTCACTTCCCTC
ACTATGGCCAACAAGTCCTCCCAGGCATCCCAGAACCGCCACTCCATGGAGATCAGCCCCCCTGTCCTCA
TCAGCTCCAGCAACCCCACTGCTGCTGCACGGATCAGCGAGCTGTCTGGGCTCTCCTGCAGTGCCCCTTC
TCAGGTTCATATAAGTACCACCGGGTTAATTGTGACCCCGCCCCCAAGCAGCCCAGTGACAACTGGCCCC
TCGTTTACTTTCCCATCAGATGTTCCCTACCAAGCTGCCCTTGGAACTTTGAATCCTCCTCTTCCACCAC
CCCCTCTCCTGGCTGCCACTGTCCTTGCCTCCACACCACCAGGCGCCACCGCCGCCGCTGCTGCTGCTGG
AATGGGACCGAGGCCCATGGCAGGATCCACTGACCAGATTGCACATTTACGGCCGCAGACTCGCCCCAGT
GTGTATGTTGCTATATATCCATACACTCCTCGGAAAGAGGATGAACTAGAGCTGAGAAAAGGGGAGATGT
TTTTAGTGTTTGAGCGCTGCCAGGATGGCTGGTTCAAAGGGACATCCATGCATACCAGCAAGATAGGGGT
TTTCCCTGGCAATTATGTGGCACCAGTCACAAGGGCGGTGACAAATGCTTCCCAAGCTAAAGTCCCTATG
TCTACAGCTGGCCAGACAAGTCGGGGAGTGACCATGGTCAGTCCTTCCACGGCAGGAGGGCCTGCCCAGA
AGCTCCAGGGAAATGGCGTGGCTGGGAGTCCCAGTGTTGTCCCCGCAGCTGTGGTATCAGCAGCTCACAT
CCAGACAAGTCCTCAGGCTAAGGTCTTGTTGCACATGACGGGGCAAATGACAGTCAACCAGGCCCGCAAT
GCTGTGAGGACAGTTGCAGCGCACAACCAGGAACGCCCCACGGCAGCAGTGACACCCATCCAGGTACAGA
ATGCCGCCGGCCTCAGCCCTGCATCTGTGGGCCTGTCCCATCACTCGCTGGCCTCCCCACAACCTGCGCC
TCTGATGCCAGGCTCAGCCACGCACACTGCTGCCATCAGTATCAGTCGAGCCAGTGCCCCTCTGGCCTGT
GCAGCAGCTGCTCCACTGACTTCCCCAAGCATCACCAGTGCTTCTCTGGAGGCTGAGCCCAGTGGCCGGA
TAGTGACCGTTCTCCCTGGACTCCCCACATCTCCTGACAGTGCTTCATCAGCTTGTGGGAACAGTTCAGC
AACCAAACCAGACAAGGATAGCAAAAAAGAAAAAAAGGGTTTGTTGAAGTTGCTTTCTGGCGCCTCCACT
AAACGGAAGCCCCGCGTGTCTCCTCCAGCATCGCCCACCCTAGAAGTGGAGCTGGGCAGTGCAGAGCTTC
CTCTCCAGGGAGCGGTGGGGCCCGAACTGCCACCAGGAGGTGGCCATGGCAGGGCAGGCTCCTGCCCTGT
GGACGGGGACGGACCGGTCACGACTGCAGTGGCAGGAGCAGCCCTGGCCCAGGATGCTTTTCATAGGAAG
GCAAGTTCCCTGGACTCCGCAGTTCCCATCGCTCCACCTCCTCGCCAGGCCTGTTCCTCCCTGGGTCCTG
TCTTGAATGAGTCTAGACCTGTCGTTTGTGAAAGGCACAGGGTGGTGGTTTCCTATCCTCCTCAGAGTGA
GGCAGAACTTGAACTTAAAGAAGGAGATATTGTGTTTGTTCATAAAAAACGAGAGGATGGCTGGTTCAAA
GGCACATTACAACGTAATGGGAAAACTGGCCTTTTCCCAGGAAGCTTTGTGGAAAACATATGAGGAGACT
GACACTGAAGAAGCTTAAAATCACTTCACACAACAAAGTAGCACAAAGCAGTTTAACAGAAAGAGCACAT
TTGTGGACTTCCAGATGGTCAGGAGATGAGCAAAGGATTGGTATGTGACTCTGATGCCCCAGCACAGTTA
CCCCAGCGAGCAGAGTGAAGAAGATGTTTGTGTGGGTTTTGTTAGTCTGGATTCGGATGTATAAGGTGTG
CCTTGTACTGTCTGATTTACTACACAGAGAAACTTTTTTTTTTTTTAAGATATATGACTAAAATGGACA
ATTGTTTACAAGGCTTAACTAATTTATTTGCTTTTTTAAACTTGAACTTTTCGTATAATAGATACGTTCT
TTGGATTATGATTTTAAGAAATTATTAATTTATGAAATGATAGGTAAGGAGAAGCTGGATTATCTCCTGT
TGAGAGCAAGAGATTCGTTTTGACATAGAGTGAATGCATTTTCCCCTCTCCTCCTCCCTGCTACCATTAT
ATTTTGGGGTTATGTTTTGCTTCTTTAAGATAGAAATCCCAGTTCTCTAATTTGGTTTTCTTCTTTGGGA
AACCAAACATACAAATGAATCAGTATCAATTAGGGCCTGGGGTAGAGAGACAGAAACTTGAGAGAAGAGA
AGTTAGTGATTCCCTCTCTTTCTAGTTTGGTAGGAATCACCCTGAAGACCTAGTCCTCAATTTAATTGTG
TGGGTTTTTAATTTTCCTAGAATGAAGTGACTGAAACAATGAGAAAGAATACAGCACAACCCTTGAACAA
AATGTATTTAGAAATATATTTAGTTTTATAGCAGAAGCAGCTCAATTGTTTGGTTGGAAAGTAGGGGAAA
TTGAAGTTGTAGTCACTGTCTGAGAATGGCTATGAAGCGTCATTTCACATTTTACCCCAACTGACCTGCA
TGCCCAGGACACAAGTAAAACATTTGTGAGATAGTGGTGGTAAGTGATGCACTCGTGTTAAGTCAAAGGC
TATAAGAAACACTGTGAAAAGTTCATATTCATCCATTGTGATTCTTTCCCCACGTCTTGCATGTATTACT
GGATTCCCACAGTAATATAGACTGTGCATGGTGTGTATATTTCATTGCGATTTCCTGTTAAGATGAGTTT
GTACTCAGAATTGACCAATTCAGGAGGTGTAAAAATAAACAGTGTTCTCTTCTCTACCCCAAAGCCACTA

-to be continued

Figure 3: Human POSH cDNA Sequence (SEQ ID NO:3)

CTGACCAAGGTCTCTTCAGTGCACTCGCTCCCTCTCTGGCTAAGGCATGCATTAGCCACTACACAAGTCA
TTAGTGAAAGTGGTCTTTTATGTCCTCCCAGCAGACAGACATCAAGGATGAGTTAACCAGGAGACTACTC
CTGTGACTGTGGAGCTCTGGAAGGCTTGGTGGGAGTGAATTTGCCCACACCTTACAATTGTGGCAGGATC
CAGAAGAGCCTGTCTTTTTATATCCATTCCTTGATGTCATTGGCCTCTCCCACCGATTTCATTACGGTGC
CACGCAGTCATGGATCTGGGTAGTCCGGAAAACAAAAGGAGGGAAGACAGCCTGGTAATGAATAAGATCC
TTACCACAGTTTTCTCATGGGAAATACATAATAAACCCTTTCATCTTTTTTTTTTTCCTTTAAGAATTAA
AACTGGGAAATAGAAACATGAACTGAAAAGTCTTGCAATGACAAGAGGTTTCATGGTCTTAAAAAGATAC
TTTATATGGTTGAAGATGAAATCATTCCTAAATTAACCTTTTTTTTAAAAAAAAACAATGTATATTATGT
TCCTGTGTGTTGAATTTAAAAAAAAAAAAATACTTTACTTGGATATTCATGTAATATATAAAGGTTTGGTG
AAATGAACTTTAGTTAGGAAAAAGCTGGCATCAGCTTTCATCTGTGTAAGTTGACACCAATGTGTCATAA
TATTCTTTATTTTGGGAAATTAGTGTATTTTATAAAAATTTTAAAAAGAAAAAAGACTACTACAGGTTAA
GATAATTTTTTTACCTGTCTTTTCTCCATATTTTAAGCTATGTGATTGAAGTACCTCTGTTCATAGTTTC
CTGGTATAAAGTTGGTTAAAATTTCATCTGTTAATAGATCATTAGGTAATATAATGTATGGGTTTTCTAT
TGGTTTTTTGCAGACAGTAGAGGGAGATTTTGTAACAAGGGCTTGTTACACAGTGATATGGTAATGATAA
AATTGCAATTTATCACTCCTTTTCATGTTAATAATTTGAGGACTGGATAAAAGGTTTCAAGATTAAAATT
TGATGTTCAAACCTTTGT

PCT/US04/06308

Figure 4: 5' cDNA fragment of human POSH (public gi:10432611; SEQ ID NO:4)

ctgagagacactgcgagcggcgagcgcggtggggccgcatctgcatcagccgccgcagccgctgcggggc
cgcgaacaaagaggaggagccgaggcgcgagagcaaagtctgaaatggatgttacatgagtcattttaag
gatgcacacaactatgaacatttctgaagatttttctcagtaaagtagataaagatggatgaatcagcc
ttgttggatcttttggagtgtccggtgtgtctagagcgccttgatgcttctgcgaaggtcttgccttgcc
agcatacgttttgcaagcgatgtttgctggggatcgtaggttctcgaaatgaactcagatgtcccgagtg
caggactcttgttggctcgggtgtcgaggagcttcccagtaacatcttgctggtcagacttctggatggc
atcaaacagaggccttggaaacctggtcctggtgggggaagtgggaccaactgcacaaatgcattaaggt
ctcagagcagcactgtggctaattgtagctcaaaagatctgcagagctcccagggcggacagcagcctcg
ggtgcaatcctggagcccccagtgaggggtatacctcagttaccatgtgccaaagcgttatacaactat
gaaggaaaagagcctggagaccttaaattcagcaaaggcgacatcatcattttgcgaagacaagtggatg
aaaattggtaccatggggaagtcaatggaatccatggcttttccccaccaactttgtgcagattattaa
accgttacctcagcccccacctcagtgcaaagcactttatgactttgaagtgaaagacaaggaagcagac
aaagattgccttccatttgcaaaggatgatgttctgactgtgatccgaagagtggatgaaaactgggctg
aaggaatgctggcagacaaaataggaatatttccaatttcatatgttgagtttaactcggctgctaagca
gctgatagaatgggataagcctcctgtgccaggagttgatgctggagaatgttcctcggcagcagcccag
agcagcactgccccaaagcactccgacaccaagaagaacaccaaaaagcggcactccttcacttccctca
ctatggccaacaagtcctcccaggcatcccagaaccgccactccatggagatcagccccctgtcctcat
cagctccagcaaccccactgctgctgcacggatcagcgagctgtctgggctctcctgcagtgccccttct
caggttcatataagtaccaccgggttaattgtgaccccgcccccaagcagcccagtgacaactggcccct
cgtttactttcccatcagatgttccctaccaagctgcccttggaactttgaatcctcctcttccaccacc
ccctctcctggctgccactgtccttgcctccacaccaccaggcgccaccgccgccgctgctgctgctgga
atgggaccgaggcccatggcaggatccactgaccagattgcacatttacggccgcagactcgccccagtg
tgtatgttgctatatatccatacactcctcggaaagaggatgaactagagctgagaaaggggagatgtt
tttagtgtttgagcgctgccaggatggctggttcaaagggacatccatgcataccagcaagataggggtt
ttccctggcaattatgtggcaccagtcacaagggcggtgacaaatgcttcccaagctaaagtccctatgt
ctacagctggccagacaagtcggggagtgaccatggtcagtccttccacggcaggagggcctgcccagaa
gctccagggaaatggcgtggctgggagtcccagtgttgtccccgcagctgtggtatcagcagctcacatc
cagacaagtcctcaggctaaggtcttgttgcacatgacggggcaaatgacagtcaaccaggcccgcaatg
ctgtgaggacagttgcagcgcacaaccaggaacgccccacggcagcagtgacacccatccaggtacagaa
tgccgccggcctcagccctgcatctgtgggcctgtcccatcactcgctggcctccccacaacctgcgcct
ctgatgccaggctcagccacgcacactgctgccatcagtatcagtcgagccagtgcccctctggcctgtg
cagcagctgctccactgacttccccaagcatcaccagtgcttctctggaggctgagcccagtggccggat
agtgaccgttctccctggactccccacatctcctgacagtgcttcatcagcttgtgggaacagttcagca
accaaaccagacaaggatagc

Figure 5: N terminus protein fragment of hPOSH (public gi:10432612; SEQ ID NO:5)

MDESALLDLLECPVCLERLDASAKVLPCQHTFCKRCLLGIVGSRNELRCPECRTLVGSGVEELPSNILLV
RLLDGIKQRPWKPGPGGGSGTNCTNALRSQSSTVANCSSKDLQSSQGGQQPRVQSWSPPVRGIPQLPCAK
ALYNYEGKEPGDLKFSKGDIIILRRQVDENWYHGEVNGIHGFFPTNFVQIIKPLPQPPPQCKALYDFEVK
DKEADKDCLPFAKDDVLTVIRRVDENWAEGMLADKIGIFPISYVEFNSAAKQLIEWDKPPVPGVDAGECS
SAAAQSSTAPKHSDTKKNTKKRHSFTSLTMANKSSQASQNRHSMEISPPVLISSSNPTAAARISELSGLS
CSAPSQVHISTTGLIVTPPPSSPVTTGPSFTFPSDVPYQAALGTLNPPLPPPPLLAATVLASTPPGATAA
AAAAGMGPRPMAGSTDQIAHLRPQTRPSVYVAIYPYTPRKEDELELRKGEMFLVFERCQDGWFKGTSMHT
SKIGVFPGNYVAPVTRAVTNASQAKVPMSTAGQTSRGVTMVSPSTAGGPAQKLQGNGVAGSPSVVPAAVV
SAAHIQTSPQAKVLLHMTGQMTVNQARNAVRTVAAHNQERPTAAVTPIQVQNAAGLSPASVGLSHHSLAS
PQPAPLMPGSATHTAAISISRASAPLACAAAAPLTSPSITSASLEAEPSGRIVTVLPGLPTSPDSASSAC
GNSSATKPDKDS

Figure 6: 3' mRNA fragment of hPOSH (public gi:7959248; SEQ ID NO:6)

atttcatatgttgagtttaactcggctgctaagcagctgatagaatgggataagcctcctgtgccaggag
ttgatgctggagaatgttcctcggcagcagcccagagcagcactgccccaaagcactccgacaccaagaa
gaacaccaaaaagcggcactccttcacttccctcactatggccaacaagtcctcccaggcatcccagaac
cgccactccatggagatcagccccctgtcctcatcagctccagcaaccccactgctgctgcacggatca
gcgagctgtctgggctctcctgcagtgccccttctcaggttcatataagtaccaccgggttaattgtgac
cccgcccccaagcagcccagtgacaactggcccctcgtttactttcccatcagatgttccctaccaagct
gcccttggaactttgaatcctcctcttccaccacccccctctcctggctgccactgtccttgcctccacac
caccaggcgccaccgccgctgctgctgctgctggaatgggaccgaggcccatggcaggatccactgacca
gattgcacatttacggccgcagactcgccccagtgtgtatgttgctatatatccatacactcctcggaaa
gaggatgaactagagctgagaaaggggagatgttttagtgtttgagcgctgccaggatggctggttca
aagggacatccatgcataccagcaagataggggttttccctggcaattatgtggcaccagtcacaagggc
ggtgacaaatgcttcccaagctaaagtccctatgtctacagctggccagacaagtcggggagtgaccatg
gtcagtccttccacggcaggagggcctgcccagaagctccagggaaatggcgtggctgggagtcccagtg
ttgtccccgcagctgtggtatcagcagctcacatccagacaagtcctcaggctaaggtcttgttgcacat
gacggggcaaatgacagtcaaccaggcccgcaatgctgtgaggacagttgcagcgcacaaccaggaacgc
cccacggcagcagtgacacccatccaggtacagaatgccgccggcctcagccctgcatctgtgggcctgt
cccatcactcgctggcctccccacaacctgcgcctctgatgccaggctcagccacgcacactgctgccat
cagtatcagtcgagccagtgcccctctggcctgtgcagcagctgctccactgacttccccaagcatcacc
agtgcttctctggaggctgagcccagtggccggatagtgaccgttctccctggactccccacatctcctg
acagtgcttcatcagcttgtgggaacagttcagcaaccaaaccagacaaggatagcaaaaagaaaaaaa
gggtttgttgaagttgctttctggcgcctccactaaacggaagccccgcgtgtctcctccagcatcgccc
accctagaagtggagctgggcagtgcagagcttcctctccagggagcggtggggcccgaactgccaccag
gaggtggccatggcagggcaggctcctgccctgtggacggggacggaccggtcacgactgcagtggcagg
agcagccctggcccaggatgcttttcataggaaggcaagttccctggactccgcagttcccatcgctcca
cctcctcgccaggcctgttcctccctgggtcctgtcttgaatgagtctagacctgtcgtttgtgaaaggc
acagggtggtggtttcctatcctcctcagagtgaggcagaacttgaacttaaagaaggagatattgtgtt
tgttcataaaaaacgagaggatggctggttcaaaggcacattacaacgtaatgggaaaactggccttttc
ccaggaagctttgtggaaaacatatgaggagactgacactgaagaagcttaaaatcacttcacacaacaa
agtagcacaaagcagtttaacagaaagagcacatttgtggacttccagatggtcaggagatgagcaaagg
attggtatgtgactctgatgccccagcacagttaccccagcgagcagagtgaagaagatgtttgtgtggg
ttttgttagtctggattcggatgtataaggtgtgccttgtactgtctgatttactacacagagaaacttt
ttttttttttaagatatatgactaaaatggacaattgtttacaaggcttaactaatttatttgcttttt
taaacttgaacttttcgtataatagatacgttctttggattatgattttaagaaattattaatttatgaa
atgataggtaaggagaagctggattatctcctgttgagagcaagagattcgttttgacatagagtgaatg
cattttcccctctcctcctccctgctaccattatattttggggttatgttttgcttctttaagatagaaa
tcccagttctctaatttggttttcttctttgggaaaccaaacatacaaatgaatcagtatcaattagggc
ctggggtagagagacagaaacttgagagaagagaagttagtgattccctctctttctagtttggtaggaa
tcaccctgaagacctagtcctcaatttaattgtgtgggtttttaattttcctagaatgaagtgactgaaa
caatgagaaagaatacagcacaacccttgaacaaaatgtatttagaaatatatttagttttatagcagaa
gcagctcaattgtttggttggaaagtaggggaaattgaagttgtagtcactgtctgagaatggctatgaa
gcgtcatttcacattttaccccaactgacctgcatgcccaggacacaagtaaaacatttgtgagatagtg
gtggtaagtgatgcactcgtgttaagtcaaaggctataagaaacactgtgaaaagttcatattcatccat
tgtgattctttccccacgtcttgcatgtattactggattcccacagtaatatagactgtgcatggtgtgt
atatttcattgcgatttcctgttaagatgagtttgtactcagaattgaccaattcaggaggtgtaaaaat
aaacagtgttctcttctctaccccaaagccactactgaccaaggtctcttcagtgcactcgctccctctc
tggctaaggcatgcattagccactacacaagtcattagtgaaagtggtcttttatgtcctcccagcagac
agacatcaaggatgagttaaccaggagactactcctgtgactgtggagctctggaaggcttggtgggagt
gaatttgcccacaccttacaattgtggcaggatccagaagagcctgtcttttatatccattccttgatg
tcattggcctctcccaccgatttcattacggtgccacgcagtcatggatctgggtagtccggaaaacaaa
aggagggaagacagcctggtaatgaataagatccttaccacagttttctcatgggaaatacataataaac
cctttcatcttttttttttttcctttaagaattaaaactgggaaatagaaacatgaactgaaaagtcttgc
aatgacaagaggtttcatggtcttaaaagatactttatatggttgaagatgaaatcattcctaaattaa
ccttttttttaaaaaaaaacaatgtatattatgttcctgtgtgttgaatttaaaaaaaaaaaatacttta
cttggatattcatgtaatatataaaggtttggtgaaatgaactttagttaggaaaaagctggcatcagct
ttcatctgtgtaagttgacaccaatgtgtcataatattctttattttgggaaattagtgtattttataaa
aattttaaaagaaaaaagactactacaggttaagataatttttttacctgtcttttctccatattttaa
gctatgtgattgaagtacctctgttcatagtttcctggtataaagttggttaaaatttcatctgttaata
gatcattaggtaatataatgtatgggttttctattggtttttttgcagacagtagagggagattttgtaac
aagggcttgttacacagtgatatggtaatgataaaattgcaatttatcactccttttcatgttaataatt
tgaggactggataaaggtttcaagattaaaatttgatgttcaaacctttgt

Figure 7: C terminus protein fragment of hPOSH (public gi:7959249; SEQ ID NO:7)

ISYVEFNSAAKQLIEWDKPPVPGVDAGECSSAAAQSSTAPKHSDTKKNTKKRHSFTSLTMANKSSQASQN
RHSMEISPPVLISSSNPTAAARISELSGLSCSAPSQVHISTTGLIVTPPPSSPVTTGPSFTFPSDVPYQA
ALGTLNPPLPPPPLLAATVLASTPPGATAAAAAAGMGPRPMAGSTDQIAHLRPQTRPSVYVAIYPYTPRK
EDELELRKGEMFLVFERCQDGWFKGTSMHTSKIGVFPGNYVAPVTRAVTNASQAKVPMSTAGQTSRGVTM
VSPSTAGGPAQKLQGNGVAGSPSVVPAAVVSAAHIQTSPQAKVLLHMTGQMTVNQARNAVRTVAAHNQER
PTAAVTPIQVQNAAGLSPASVGLSHHSLASPQPAPLMPGSATHTAAISISRASAPLACAAAAPLTSPSIT
SASLEAEPSGRIVTVLPGLPTSPDSASSACGNSSATKPDKDSKKEKKGLLKLLSGASTKRKPRVSPPASP
TLEVELGSAELPLQGAVGPELPPGGGHGRAGSCPVDGDGPVTTAVAGAALAQDAFHRKASSLDSAVPIAP
PPRQACSSLGPVLNESRPVVCERHRVVVSYPPQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKTGLF
PGSFVENI

Figure 8: Human POSH full mRNA, Annotated Sequence (part 1)

```
---- - gi|10432611|dbj|AK021429.1|AK021429 Homo sapiens cDNA,
FLJ11367 fis, clone HEMBA1000303, highly similar to Mus musculus
Plenty of SH3s (POSH) mRNA

---- - gi|7959248|dbj|AB040927.1|AB040927 Homo sapiens mRNA for
KIAA1494 protein, partial cds

- Both hPOSH and KIAA1495

- Ring Domain

- SH3 Domian

- start codon and stop codon of predicted ORF
```

```
CTGAGAGACACTGCGAGCGGCGAGCGCGGTGGGGCCGCATCTGCATCAGCCGCCGCAGCCGCTGCGGGGC
CGCGAACAAAGAGGAGGAGCCGAGGCGCGAGAGCAAAGTCTGAAATGGATGTTACATGAGTCATTTTAAG
GGATGCACACAACTATGAACATTTCTGAAGATTTTTCTCAGTAAAGTAGATAAAGATGGATGAATCAGC
CTTGTTGGATCTTTTGGAG[illegible]
[illegible]
GCAGGACTCTTGTTGGCTCGGGTGTCGAGGAGCTTCCCAGTAACATCTTGCTGGTCAGACTTCTGGATGG
CATCAAACAGAGGCCTTGGAAACCTGGTCCTGGTGGGGAAGTGGGACCAACTGCACAAATGCATTAAGG
TCTCAGAGCAGCACTGTGGCTAATTGTAGCTCAAAAGATCTGCAGAGCTCCCAGGGCGGACAGCAGCCTC
GGGTGCAATCCTGGAGCCCCCCAGTGAGGGGTATACCTCAGTTA[illegible]
[illegible]
[illegible]A
AACCGTTACCTCAGCCCCCA[illegible]
[illegible]
[illegible]TCGCCTCCTAACG
AGCTGATAGAATGCCATAAGCCTCCTGTGCCACGAGTTCATCCTGCACAATGTTCCTCGGCTACCAGCCCA
GAGCAGCAGTGCCCCAAAGCAGTCCCACACCAAGAAGAACACCAAAAAGCGGCACTCCTTGACTTCCGTC
ACTATGCCCAACAAGTCCTCCGAGGCATCCCAGAACCGGCACTCCATGCACATCAGCCGCCCTGTCCTCA
TCAGCTCCACCAACCCCACTGCTGCTGCACGGATCAGGCAGCTGTGTCGGCTCTCCTGCAGTGCCCCTTG
TCAGGTTCATATAAGTACCACCGCCGTTAATTCTGACCCGCCGCCCCAAGCACCCCAGTCAGAACTGGCCCC
TCGTTTACTTTCCCATCAGATCTTGCCTACCAAGCTCCCCTTGCAACTTTGAATCCTCCTCTTCCACCAC
CCCCTCTCCTGCCTCCCACTGTCCTTGCCTCGACACCACCAGGCCGCGACCCGCCGCCCCTGCTCCTGCTCC
AATGCCACCGCAGCCGCCATCGCAGCATCCACTCACCCAGATTGCACATTTACGGCCGCAGACTCCCCC[illegible]
[illegible]
[illegible]
[illegible]ACAACGGCGGCTGACAAATGCTTCCCAACCTAAAGTCCCTATG
TCTAGAGCTGCCCAGAGAACTCGCCGCAGTCACCATGGTCAGTCCTTCCACCGCAGGAGCGCCTGCCCACA
AGCTCCACGGAAATGCCGTGCCTCGCACTCCCAGTCTTGTCCCCGGAGCTCTCCTATCAGCAGCTCACAT
CGACACAAGTCCTGACGCTAAGCTCTTGTTCCACATCACCGGCCAAATGACAGTCAACCATGCCCGCAAT
GCTGCTGACGACAGTTGCACCCCACAACCAGGAACCGCCCCACGGCAGCAGTGAGACCCATCCAGGTAGA
ATGCCGCCCCGCTCAGCCGTCCATCTGTGCCGCTCTCCCATCACTCCGCTGCCCTGCCCAGAACCTGCCGC
TCTGATGCCACCCTCAGCCCAGCCAGACTCCCTGCCATCACTATCAGTCGAGCCAGTGCCCCCTCTGGCCTCT
GCAGCCAGCCTCCTCCACTCACTTCCGCAAGCATCACCACTGCTTCTCTGCAGCCTCAGCCCAGTGCCCGGA
TACTGACCCTTCTCCCGTGCACTCCCGCACATCTGCTCACACTGCTTCATCAGCTTCTGGCAAGAGTTCAGC
AACCAAACCAGACAAGCGATAGCAAAAAAGAAAAAAAGGGTTTGTTGAAGTTGCTTTCTGGCGCCTCCACT
AAACGGAAGCCCCGCGTGTCTCCTCCAGCATCGCCCACCCTAGAAGTGGAGCTGGGCAGTGCAGAGCTTC
CTCTCCAGGGAGCGGTGGGGCCCGAACTGCCACCAGGAGGTGGCCATGGCAGGGCAGGCTCCTGCCCTGT
GGACGGGGACGGACCGGTCACGACTGCAGTGGCAGGAGCAGCCCTGGCCCAGGATGCTTTTCATAGGAAG
GCAAGTTCCCTGGACTCCGCAGTTCCCATCGCTCCACCTCCTCGCCAGGCCTGTTCCTCCCTGGGTCCTG
TCTTGAATGAGTCTAGACCTGTCGTTTGT[illegible]
[illegible]
[illegible]TGAGGAGACT
GACACTGAAGAAGCTTAAAATCACTTCACACAACAAAGTAGCACAAAGCAGTTTAACAGAAAGAGCACAT
```

-to be continued

Figure 8: Human POSH full mRNA, Annotated Sequence (part 2)

TTGTGGACTTCCAGATGGTCAGGAGATGAGCAAAGGATTGGTATGTGACTCTGATGCCCCAGCACAGTTA
CCCCAGCGAGCAGAGTGAAGAAGATGTTTGTGTGGGTTTTGTTAGTCTGGATTCGGATGTATAAGGTGTG
CCTTGTACTGTCTGATTTACTACACAGAGAAACTTTTTTTTTTTTTTAAGATATATGACTAAAATGGACA
ATTGTTTACAAGGCTTAACTAATTTATTTGCTTTTTTAAACTTGAACTTTTCGTATAATAGATACGTTCT
TTGGATTATGATTTTAAGAAATTATTAATTTATGAAATGATAGGTAAGGAGAAGCTGGATTATCTCCTGT
TGAGAGCAAGAGATTCGTTTTGACATAGAGTGAATGCATTTTCCCCTCTCCTCCTCCCTGCTACCATTAT
ATTTTGGGGTTATGTTTTGCTTCTTTAAGATAGAAATCCCAGTTCTCTAATTTGGTTTTCTTCTTTGGGA
AACCAAACATACAAATGAATCAGTATCAATTAGGGCCTGGGGTAGAGAGACAGAAACTTGAGAGAAGAGA
AGTTAGTGATTCCCTCTCTTTCTAGTTTGGTAGGAATCACCCTGAAGACCTAGTCCTCAATTTAATTGTG
TGGGTTTTTAATTTTCCTAGAATGAAGTGACTGAAACAATGAGAAAGAATACAGCACAACCCTTGAACAA
AATGTATTTAGAAATATATTTAGTTTTATAGCAGAAGCAGCTCAATTGTTTGGTTGGAAAGTAGGGGAAA
TTGAAGTTGTAGTCACTGTCTGAGAATGGCTATGAAGCGTCATTTCACATTTTACCCCAACTGACCTGCA
TGCCCAGGACACAAGTAAAACATTTGTGAGATAGTGGTGGTAAGTGATGCACTCGTGTTAAGTCAAAGGC
TATAAGAAACACTGTGAAAAGTTCATATTCATCCATTGTGATTCTTTCCCCACGTCTTGCATGTATTACT
GGATTCCCACAGTAATATAGACTGTGCATGGTGTGTATATTTCATTGCGATTTCCTGTTAAGATGAGTTT
GTACTCAGAATTGACCAATTCAGGAGGTGTAAAAATAAACAGTGTTCTCTTCTCTACCCCAAAGCCACTA
CTGACCAAGGTCTCTTCAGTGCACTCGCTCCCTCTCTGGCTAAGGCATGCATTAGCCACTACACAAGTCA
TTAGTGAAAGTGGTCTTTTATGTCCTCCCAGCAGACAGACATCAAGGATGAGTTAACCAGGAGACTACTC
CTGTGACTGTGGAGCTCTGGAAGGCTTGGTGGAGTGAATTTGCCCACACCTTACAATTGTGGCAGGATC
CAGAAGAGCCTGTCTTTTTATATCCATTCCTTGATGTCATTGGCCTCTCCCACCGATTTCATTACGGTGC
CACGCAGTCATGGATCTGGGTAGTCCGGAAAACAAAAGGAGGGAAGACAGCCTGGTAATGAATAAGATCC
TTACCACAGTTTTCTCATGGGAAATACATAATAAACCCTTTCATCTTTTTTTTTTTCCTTTAAGAATTAA
AACTGGGAAATAGAAACATGAACTGAAAAGTCTTGCAATGACAAGAGGTTTCATGGTCTTAAAAAGATAC
TTTATATGGTTGAAGATGAAATCATTCCTAAATTAACCTTTTTTTTAAAAAAAAACAATGTATATTATGT
TCCTGTGTGTTGAATTTAAAAAAAAAAAATACTTTACTTGGATATTCATGTAATATATAAAGGTTTGGTG
AAATGAACTTTAGTTAGGAAAAAGCTGGCATCAGCTTTCATCTGTGTAAGTTGACACCAATGTGTCATAA
TATTCTTTATTTTGGGAAATTAGTGTATTTTATAAAAATTTTAAAAAGAAAAAAGACTACTACAGGTTAA
GATAATTTTTTTACCTGTCTTTTCTCCATATTTTAAGCTATGTGATTGAAGTACCTCTGTTCATAGTTTC
CTGGTATAAAGTTGGTTAAAATTTCATCTGTTAATAGATCATTAGGTAATATAATGTATGGGTTTTCTAT
TGGTTTTTTGCAGACAGTAGAGGGAGATTTTGTAACAAGGGCTTGTTACACAGTGATATGGTAATGATAA
AATTGCAATTTATCACTCCTTTTCATGTTAATAATTTGAGGACTGGATAAAAGGTTTCAAGATTAAAATT-
TGATGTTCAAACCTTTGT

Figure 9: Domain Analysis of Human POSH

Domain Name	begin	end	E-value
RING	12	52	1.06e-08
SH3	137	192	2.76e-19
SH3	199	258	4.84e-15
low complexity	366	384	-
low complexity	390	434	-
SH3	448	505	2.40e-19
low complexity	547	563	-
low complexity	652	668	-
low complexity	705	729	-
SH3	832	888	1.47e-14

PCT/US04/06308

Figure 10: Diagram of Human POSH Nucleic Acids

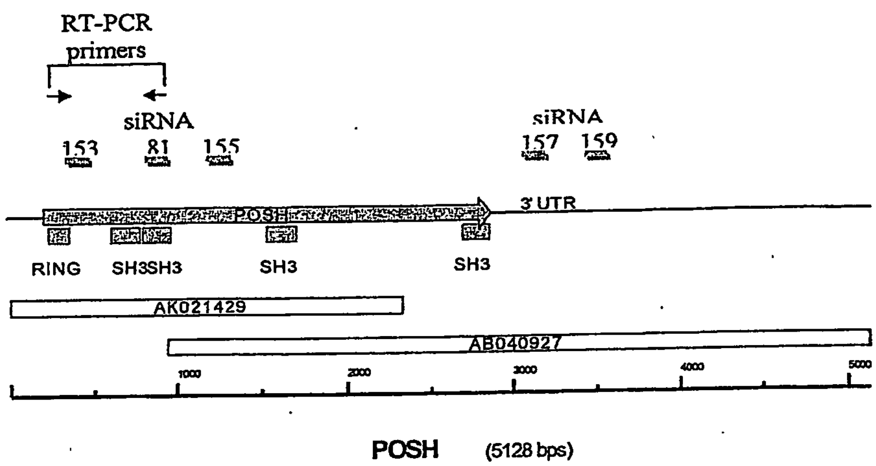


Figure 11: Reduction in Full Length POSH mRNA by siRNA Duplexes

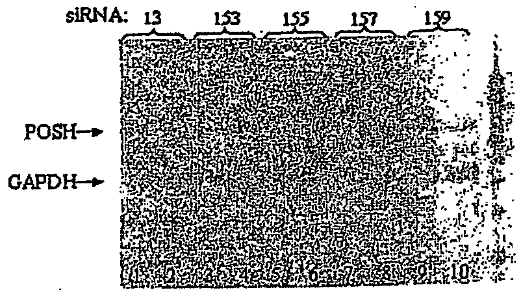


Figure 12: POSH Affects Release of VLP from Cells

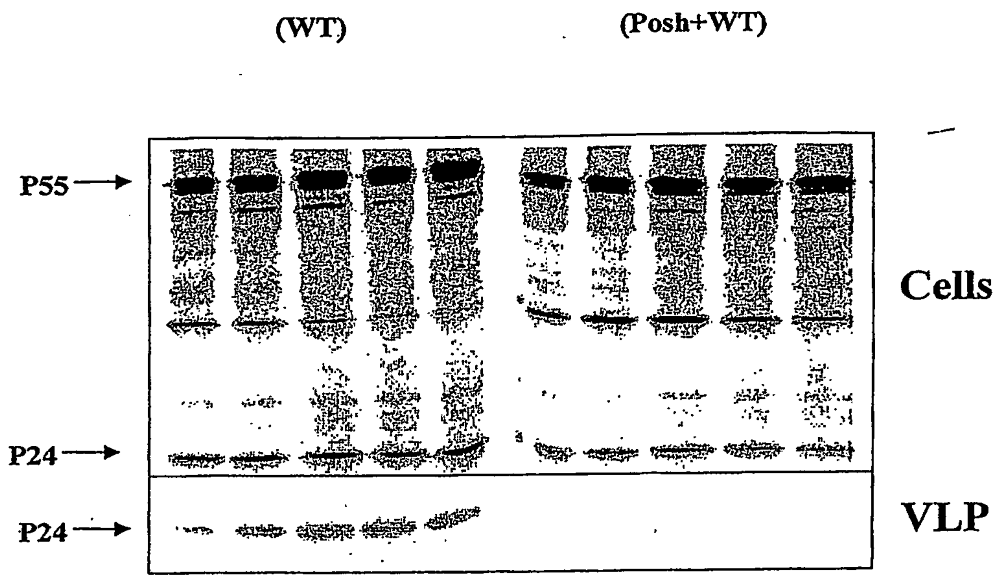


Figure 13: Release of VLP from Cells at Steady State

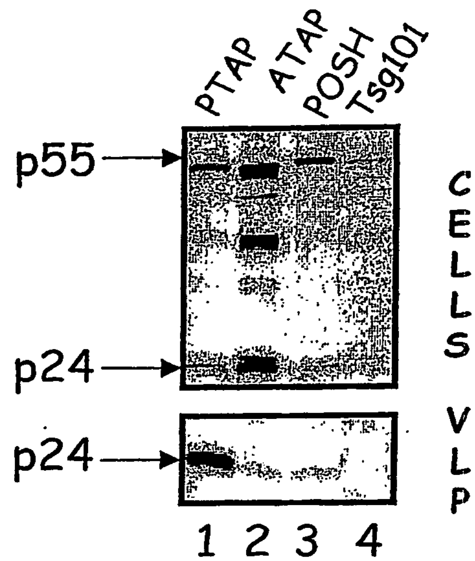


Figure 14: Mouse POSH mRNA sequence (public gi:10946921; SEQ ID NO: 8)

GGGCAGCGGGCTCGGCGGGGCTGCATCTACCAGCGCTGCGGGGCCGCGAACAAAGGCGAGCAGCGGAGGC
GCGAGAGCAAAGTCTGAAATGGATGTTACATGAATCACTTTAAGGGCTGCGCACAACTATGAACGTTCTG
AAGCCGTTTTCTCACTAAAGTCACTCAAGATGGATGAGTCTGCCTTGTTGGACCTTCTGGAGTGCCCTGT
GTGTCTAGAACGCCTGGATGCTTCCGCAAAGGTCTTACCCTGCCAGCATACCTTTTGCAAACGCTGTTTG
CTGGGGATTGTGGGTTCCCGGAATGAACTCAGATGTCCCGAATGCCGGACTCTTGTTGGCTCTGGGGTCG
ACGAGCTCCCCAGTAACATCCTACTGGTCAGACTTCTGGATGGCATCAAGCAGAGGCCTTGGAAACCCGG
CCCTGGTGGGGGCGGCGGGACCACCTGCACAAACACATTAAGGGCGCAGGGCAGCACTGTGGTTAATTGT
GGCTCGAAAGATCTGCAGAGCTCCCAGTGTGGACAGCAGCCTCGGGTGCAAGCCTGGAGCCCCCCAGTGA
GGGGAATACCTCAGTTACCGTGTGCCAAAGCATTATATAACTACGAAGGAAAAGAGCCCGGAGACCTTAA
GTTCAGCAAAGGCGACACCATCATTCTGCGCCGACAGGTGGATGAGAATTGGTACCACGGGGAAGTCAGC
GGGGTCCACGGCTTTTTCCCCACTAACTTCGTGCAGATCATCAAACCTTTACCTCAGCCCCCGCCTCAGT
GCAAAGCACTTTACGACTTTGAAGTGAAAGACAAGGAAGCTGACAAAGATTGCCTTCCCTTCGCAAAGGA
CGACGTACTGACCGTGATCCGCAGAGTGGATGAAAACTGGGCTGAAGGAATGCTGGCAGATAAAATAGGA
ATATTTCCAATTTCATACGTGGAGTTTAACTCAGCTGCCAAGCAGCTGATAGAGTGGGATAAGCCTCCCG
TGCCAGGAGTGGACACGGCAGAATGCCCCTCAGCGACGGCGCAGAGCACCTCTGCCTCAAAGCACCCCGA
CACCAAGAAGAACACCAGGAAGCGACACTCCTTCACCTCCCTCACCATGGCCAACAAGTCTTCCCAGGGG
TCCCAGAACCGCCACTCCATGGAGATCAGCCCTCCTGTGCTCATCAGTTCCAGCAACCCCACAGCCGCAG
CCCGCATCAGCGAACTGTCCGGGCTCTCCTGCAGCGCCCCGTCTCAGGTCCATATAAGCACCACTGGGTT
AATTGTGACCCCACCCCCTAGCAGCCCGGTGACAACTGGCCCTGCGTTCACGTTCCCTTCAGATGTCCCC
TACCAAGCTGCCCTTGGAAGTATGAATCCTCCACTTCCCCCACCCCCTCTCCTGGCGGCCACCGTACTCG
CCTCCACCCCGTCAGGCGCTACTGCTGCTGTTGCTGCTGCTGCTGCCGCCGCCGCCGCTGCTGGAATGGG
ACCCAGGCCTGTGATGGGGTCCTCTGAACAGATTGCACATTTACGGCCTCAGACTCGTCCCAGTGTATAT
GTTGCTATATATCCGTACACTCCCCGGAAGGAAGACGAACTGGAGCTGAGGAAAGGGGAGATGTTTTTGG
TGTTTGAGCGTTGCCAGGACGGCTGGTACAAAGGGACATCGATGCATACCAGCAAGATAGGCGTTTTCCC
TGGCAACTATGTGGCGCCCGTCACAAGGGCGGTGACGAATGCCTCCCAAGCTAAAGTCTCTATGTCTACT
GCGGGTCAGGCAAGTCGCGGGGTGACCATGGTCAGCCCTTCCACTGCAGGAGGACCTACACAGAAGCCCC
AAGGAAACGGCGTGGCCGGAAATCCCAGCGTCGTCCCCACGGCTGTGGTGTCAGCAGCTCATATCCAGAC
AAGTCCTCAGGCTAAGGTCCTGCTGCACATGTCTGGGCAGATGACAGTCAATCAGGCCCGCAATGCTGTG
AGGACAGTTGCAGCACATAGCCAGGAACGCCCCACAGCAGCAGTGACTCCCATCCAGGTCCAGAATGCCG
CCTGCCTTGGTCCTGCATCCGTGGGCCTGCCCCATCATTCTCTGGCCTCCCAACCTCTGCCTCCAATGGC
GGGTCCTGCTGCCCACGGTGCTGCCGTCAGCATCAGTCGAACCAATGCCCCCATGGCCTGCGCTGCAGGG
GCTTCTCTGGCCTCCCCAAATATGACCAGTGCCATGTTGGAGACAGAGCCCAGTGGTCGCACAGTGACCA
TCCTCCCTGGACTCCCCACATCTCCAGAGAGTGCTGCATCAGCGTGTGGGAACAGTTCAGCTGGGAAACC
AGACAAGGACAGTAAGAAAGAAAAAAAGGGCCTACTGAAGCTGCTTTCTGGTGCCTCCACCAAACGCAAG
CCCCGAGTCTCCCCTCCAGCATCACCTACCCTGGATGTGGAGCTGGGTGCTGGGGAGGCTCCCTTGCAGG
GAGCAGTAGGTCCTGAGCTGCCGCTAGGGGGCAGCCACGGCAGAGTGGGGTCATGCCCCACAGATGGTGA
TGGTCCAGTGGCCGCTGGAACAGCAGCCCTAGCCCAGGATGCCTTCCACCGCAAGACAAGCTCCCTGGAC
TCCGCAGTGCCCATTGCTCCACCACCTCGCCAGGCCTGCTCCTCCCTGGGCCCAGTCATGAATGAGGCCC
GGCCTGTTGTTTGTGAAAGGCACAGGGTGGTGGTTTCCTACCCTCCTCAGAGTGAGGCCGAACTTGAACT
CAAGGAAGGAGATATTGTGTTTGTTCATAAGAAACGAGAGGACGGCTGGTTCAAAGGCACGTTACAGAGG
AATGGGAAGACTGGCCTTTTCCCAGGGAGCTTTGTGGAAAACATCTGAGAAGACGGACACGGAGAAAGC
TTATCATCACACCACGTGTGACTAAAGAGCACAAAGCAGTTTCATAGAAAGAGCACATCTGTGGACTTCC
AGATCTTCAAGAACCGAGCAGAAGATGGGCACCTGACTCCAGAGCCCCGGCCTGGTTACCCCAGGGGCAG
AGGGAAGGAGGACACACCTGTGTGGGTTCCGTCTCTCTGGGTTCTGATGTGTAAAGTGTGCCTTGTAATG
TCTAATGGACTTTACAGATAAATGTCTTTTTTTTTTTAAGATGTATAACTAAAATGGACAATTGTTTACA
AGGCTTAACTAATTTATTTGCTTTTTTAAAACTTGAACTTTCTTGTAATAGCAAAT

Figure 15: Mouse POSH Protein sequence (Public gi: 10946922; SEQ ID NO: 9)

MDESALLDLLECPVCLERLDASAKVLPCQHTFCKRCLLGIVGSRNELRCPECRTLVGSGVDELPSNILLV
RLLDGIKQRPWKPGPGGGGGTTCTNTLRAQGSTVVNCGSKDLQSSQCGQQPRVQAWSPPVRGIPQLPCAK
ALYNYEGKEPGDLKFSKGDTIILRRQVDENWYHGEVSGVHGFFPTNFVQIIKPLPQPPPQCKALYDFEVK
DKEADKDCLPFAKDDVLTVIRRVDENWAEGMLADKIGIFPISYVEFNSAAKQLIEWDKPPVPGVDTAECP
SATAQSTSASKHPDTKKNTRKRHSFTSLTMANKSSQGSQNRHSMEISPPVLISSSNPTAAARISELSGLS
CSAPSQVHISTTGLIVTPPPSSPVTTGPAFTFPSDVPYQAALGSMNPPLPPPPLLAATVLASTPSGATAA
VAAAAAAAAAAGMGPRPVMGSSEQIAHLRPQTRPSVYVAIYPYTPRKEDELELRKGEMFLVFERCQDGWY
KGTSMHTSKIGVFPGNYVAPVTRAVTNASQAKVSMSTAGQASRGVTMVSPSTAGGPTQKPQGNGVAGNPS
VVPTAVVSAAHIQTSPQAKVLLHMSGQMTVNQARNAVRTVAAHSQERPTAAVTPIQVQNAACLGPASVGL
PHHSLASQPLPPMAGPAAHGAAVSISRTNAPMACAAGASLASPNMTSAMLETEPSGRTVTILPGLPTSPE
SAASACGNSSAGKPDKDSKKEKKGLLKLLSGASTKRKPRVSPPASPTLDVELGAGEAPLQGAVGPELPLG
GSHGRVGSCPTDGDGPVAAGTAALAQDAFHRKTSSLDSAVPIAPPPRQACSSLGPVMNEARPVVCERHRV
VVSYPPQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKTGLFPGSFVENI

Figure 16: Drosophila melanogaster POSH mRNA sequence (public gi:17737480; SEQ ID NO:10)

CATTTGTATCCGCTTGGCCACGAGCTTTGGCTGCACTTGGCAAACTTAATAAATTAAACATTGAATCCTG
CCTATTGCAACGATAATATAATCTGATTTAGTGCATTAAGAACGACAAGTAGCGATTATAATAGTAGATT
TTAGCATTTGAGCTAAATTTATTTCCCAACCGCGTCTTGGGATTGCGTATGCGTGAGCCAGTACCTGCAT
GTGTGTGTGTTTTGGAATGTGGCCCTGCACGAAATTCAAATAGTGACCATCCTTGAGATTTTGCATACTG
GCAAGATGGACGAGCACACGTTAAACGACCTGTTGGAGTGCTCCGTGTGTCTTGAGCGACTGGACACCAC
ATCGAAGGTGCTGCCATGCCAGCACACCTTCTGCCGCAAATGCTTGCAGGACATTGTGGCCAGTCAGCAC
AAGTTGCGATGCCCGGAGTGCCGCATCCTGGTCTCTTGCAAAATTGATGAGCTGCCTCCAAACGTCTTGC
TGATGCGAATCTTAGAAGGCATGAAACAAAATGCAGCAGCTGGCAAAGGAGAAGAAAAGGGAGAGGAGAC
TGAAACACAGCCGGAAAGGGCCAAACCTCAGCCGCCAGCGGAATCAGTGGCCCCGCCTGACAACCAACTA
CTCCAGCTGCAGTCACATCAGCAATCTCATCAGCCGGCTCGTCACAAGCAACGTCGATTTCTACTCCCCC
ACGCCTATGCCCTCTTTGACTTCGCCTCCGGTGAAGCCACCGATCTAAAGTTCAAGAAAGGGGATCTGAT
ACTGATCAAGCATCGCATCGACAACAACTGGTTTGTGGGTCAAGCGAATGGTCAGGAGGGCACATTTCCC
ATCAACTACGTCAAGGTATCGGTTCCGCTGCCCATGCCGCAGTGCATTGCCATGTATGACTTTAAGATGG
GGCCCAACGACGAGGAGGGATGCCTCGAATTTAAGAAAAGCACTGTAATACAGGTAATGCGCCGAGTTGA
TCATAATTGGGCAGAAGGACGAATTGGCCAGACCATCGGAATCTTTCCAATAGCATTCGTTGAGCTGAAT
GCAGCGGCCAAAAAGCTGTTGGACAGCGGGCTACACACCCATCCATTCTGCCATCCACCGAAGCAACAGG
GGCAGCGGGCCCTTCCTCCGGTTCCAGTTATTGATCCCACGGTGGTCACGGAATCCAGTTCGGGATCCTC
CAATTCCACGCCGGGCAGCAGCAATTCAAGCTCCACATCCAGCTCGAATAACTGCAGTCCGAATCACCAA
ATCTCACTGCCGAATACCCCCCAACATGTAGTAGCTTCCGGATCGGCGTCTGTTCGTTTCCGTGACAAGG
GAGCAAAGGAGAAACGCCACTCACTAAATGCTTTGCTGGGAGGAGGAGCTCCATTAAGTCTGCTGCAGAC
CAACCGCCATTCGGCTGAAATTCTTAGCCTGCCCCATGAACTAAGCCGCTTGGAAGTTTCCAGCTCAACA
GCTCTAAAACCCACGTCAGCCCCACAGACATCGCGTGTACTTAAGACCACTGTTCAGCAGCAGATGCAAC
CGAATTTACCCTGGGGATACTTAGCCCTGTTCCCATACAAACCACGCCAAACGGATGAGCTGGAATTAAA
AAAGGGTTGTGTTTACATTGTGACCGAACGATGTGTGGACGGTTGGTTCAAGGGAAAAAACTGGTTGGAC
ATCACTGGAGTGTTCCCGGGCAACTACCTGACGCCCCTGCGCGCCCGCGACCAGCAGCAGTTAATGCATC
AATGGAAATATGTTCCCCAAAATGCAGACGCCCAGATGGCACAAGTACAGCAGCATCCAGTTGCACCAGA
TGTGCGACTCAACAACATGCTGTCCATGCAACCGCCTGATTTGCCACCTCGTCAGCAGCAGGCTACCGCC
ACGACCACCAGTTGCTCTGTGTGGTCGAAACCAGTGGAGGCGCTGTTCAGCAGAAAATCGGAGCCCAAGC
CTGAAACTGCCACAGCTTCGACTACGAGCAGCAGTTCCTCTGGAGCAGTGGGACTTATGAGGAGATTAAC
TCACATGAAAACACGCTCCAAATCTCCGGGAGCGTCCTTGCAGCAAGTTCCGAAAGAAGCTATTAGCACA
AATGTGGAATTTACAACAAACCCATCAGCTAAATTGCATCCAGTACATGTAAGATCCGGCTCGTGCCCCA
GTCAGCTGCAGCACAGTCAACCGCTCAATGAAACTCCAGCAGCCAAGACAGCGGCACAACAACAGCAGTT
CCTACCCAAGCAGCTGCCTTCCGCTTCTACGAACAGCGTTTCGTACGGATCGCAACGCGTGAAAGGAAGC
AAGGAACGTCCTCACTTGATTTGCGCGAGACAATCATTAGATGCAGCTACATTTCGCAGTATGTACAACA
ATGCCGCGTCGCCGCCGCCACCTACTACTTCCGTGGCCCCAGCTGTCTACGCCGGCGGTCAGCAACAGGT
GATTCCTGGAGGTGGAGCGCAATCCCAGTTGCATGCCAATATGATTATTGCACCCAGCCATCGGAAGTCG
CACAGCCTAGATGCGAGTCATGTGCTGAGTCCCAGCAGCAATATGATCACGGAGGCGGCCATTAAGGCCA
GCGCCACCACTAAGTCTCCTTACTGCACGAGGGAAAGTCGATTCCGCTGCATTGTGCCGTATCCACCAAA
CAGTGACATTGAACTAGAGCTACATTTGGGCGACATTATCTACGTCCAGCGGAAGCAGAAGAACGGCTGG
TATAAGGGCACCCATGCCCGTACCCACAAAACCGGGCTGTTCCCCGCCTCCTTTGTTGAACCGGATTGTT
AGGAAAGTTATGGTTCAAACTAGAATTTATTAAGCGAAATTCCAAATTACTTGTCTAAAAGGATTCAATC
GTCGGTCTATTCGGGCTTCCAAATACGCAATCTCATATTTCTCTTTTCAAAAAAGAAACCGTTTTGTACT
CTTCCAATCGAATGGGCAGCTCGCCGTTGTACTTTTTTATACAATGCTTGATCAAAATAGGCTAGCCATG
TAAGACTTAGGGAACAGTTACTTAAGCCTTAGCGATTAGTTAGCTAGAGAAATAATCTAACCGATCCTTG
TGCCCTCTACAAAGTTATTTGTAATATACGATACTCAGTAATAAAAAAAAAAAAAAAAAAAAAAAAAA

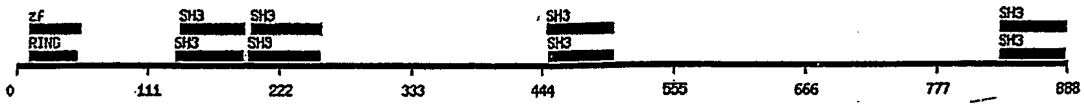
Figure 17: Drosophila melanogaster POSH protein sequence (public gi:17737481; SEQ ID NO:11)

MDEHTLNDLLECSVCLERLDTTSKVLPCQHTFCRKCLQDIVASQHKLRCPECRILVSCKIDELPPNVLLM
RILEGMKQNAAAGKGEEKGEETETQPERAKPQPPAESVAPPDNQLLQLQSHQQSHQPARHKQRRFLLPHA
YALFDFASGEATDLKFKKGDLILIKHRIDNNWFVGQANGQEGTFPINYVKVSVPLPMPQCIAMYDFKMGP
NDEEGCLEFKKSTVIQVMRRVDHNWAEGRIGQTIGIFPIAFVELNAAAKKLLDSGLHTHPFCHPPKQQGQ
RALPPVPVIDPTVVTESSSGSSNSTPGSSNSSSTSSSNNCSPNHQISLPNTPQHVVASGSASVRFRDKGA
KEKRHSLNALLGGGAPLSLLQTNRHSAEILSLPHELSRLEVSSSTALKPTSAPQTSRVLKTTVQQQMQPN
LPWGYLALFPYKPRQTDELELKKGCVYIVTERCVDGWFKGKNWLDITGVFPGNYLTPLRARDQQQLMHQW
KYVPQNADAQMAQVQQHPVAPDVRLNNMLSMQPPDLPPRQQQATATTTSCSVWSKPVEALFSRKSEPKPE
TATASTTSSSSSGAVGLMRRLTHMKTRSKSPGASLQQVPKEAISTNVEFTTNPSAKLHPVHVRSGSCPSQ
LQHSQPLNETPAAKTAAQQQQFLPKQLPSASTNSVSYGSQRVKGSKERPHLICARQSLDAATFRSMYNNA
ASPPPPTTSVAPAVYAGGQQQVIPGGGAQSQLHANMIIAPSHRKSHSLDASHVLSPSSNMITEAAIKASA
TTKSPYCTRESRFRCIVPYPPNSDIELELHLGDIIYVQRKQKNGWYKGTHARTHKTGLFPASFVEPDC

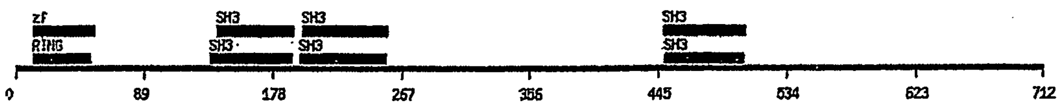
PCT/US04/06308

Figure 18: POSH Domain Analysis

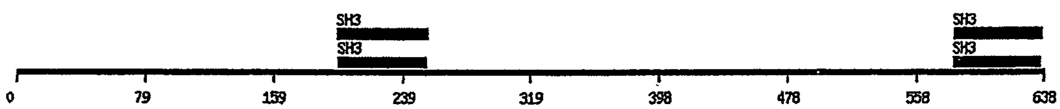
hPOSH protein sequence :



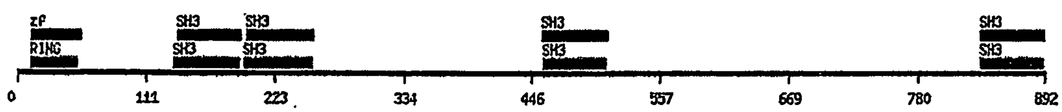
N terminus protein fragment of hPOSH (public gi:10432612):



C terminus protein fragment of hPOSH (public gi:7959249):

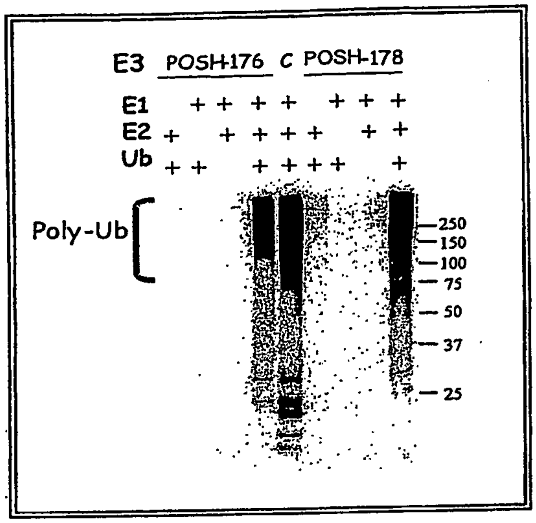


Mouse POSH Protein sequence (Public gi: 10946922):



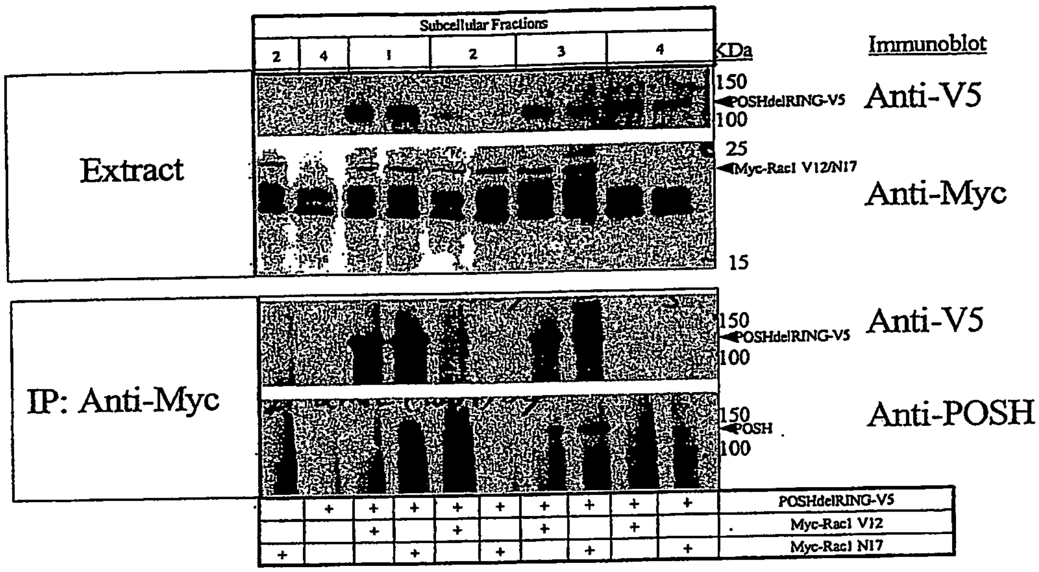
Drosophila melanogaster POSH protein sequence (public gi:17737481)

Figure 19: Human POSH has ubiquitin ligase activity



PCT/US04/06308

Figure 20



PCT/US04/06308

Figure 21. PLD activity in medium of transfected cells

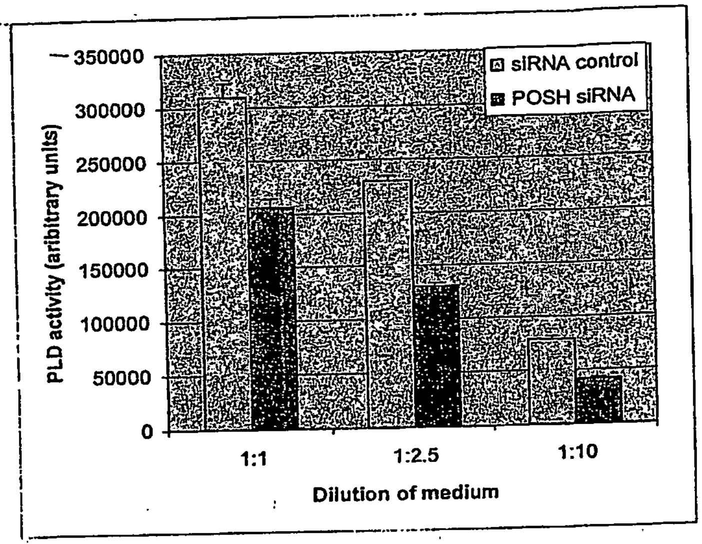


Figure 22.

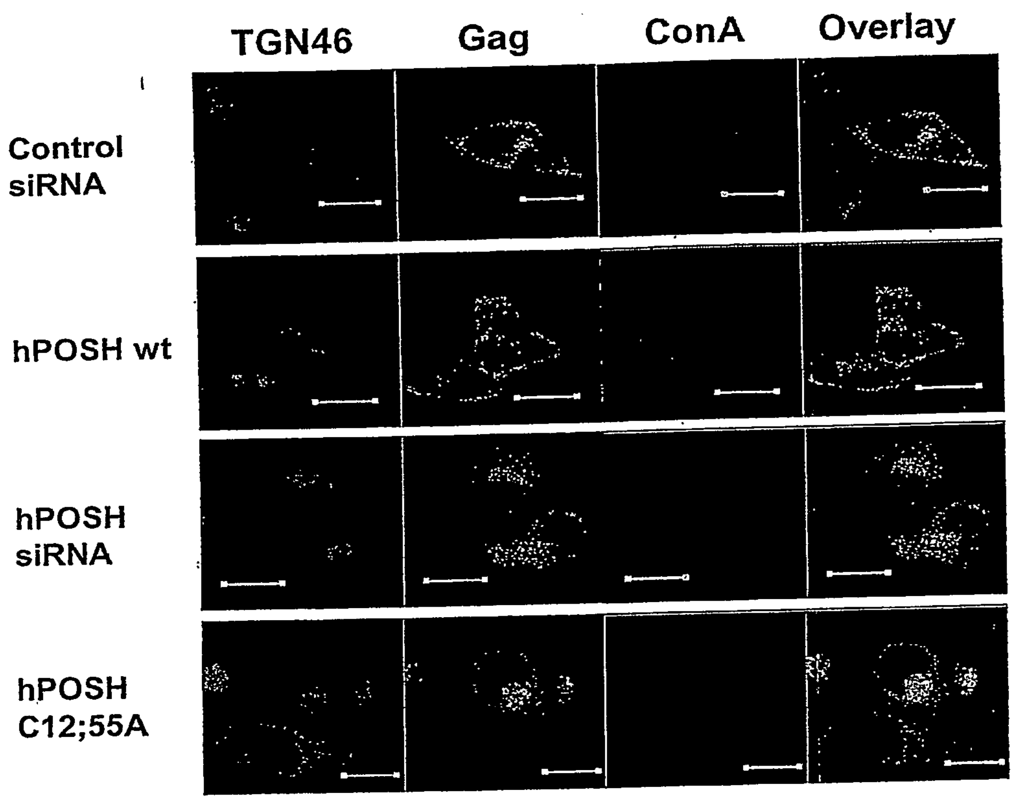
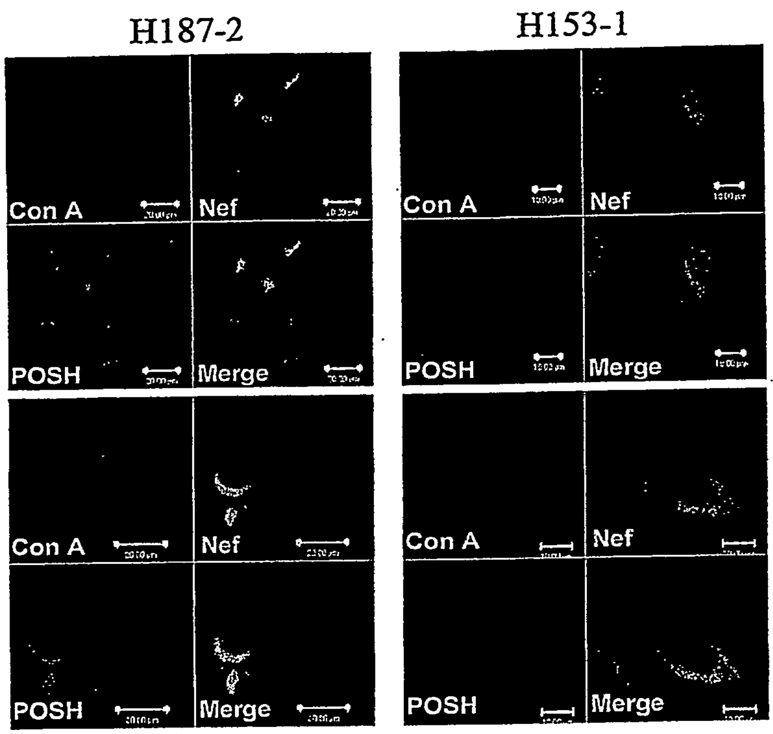

TGN46
Gag
ConA
Overlay
Control siRNA
hPOSH wt
hPOSH siRNA
hPOSH C12;55A

Figure 23.


H187-2
H153-1
Con A
Nef
POSH
Merge
Con A
Nef
POSH
Merge
Con A
Nef
POSH
Merge
Con A
Nef
POSH
Merge

5

PCT/US04/06308

Figure 24.

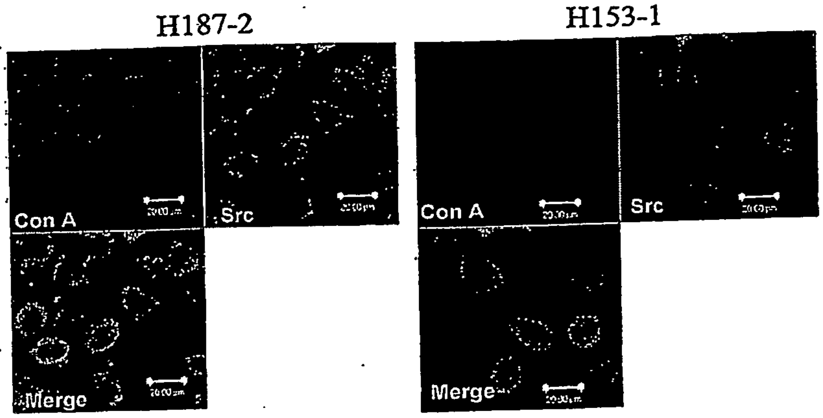
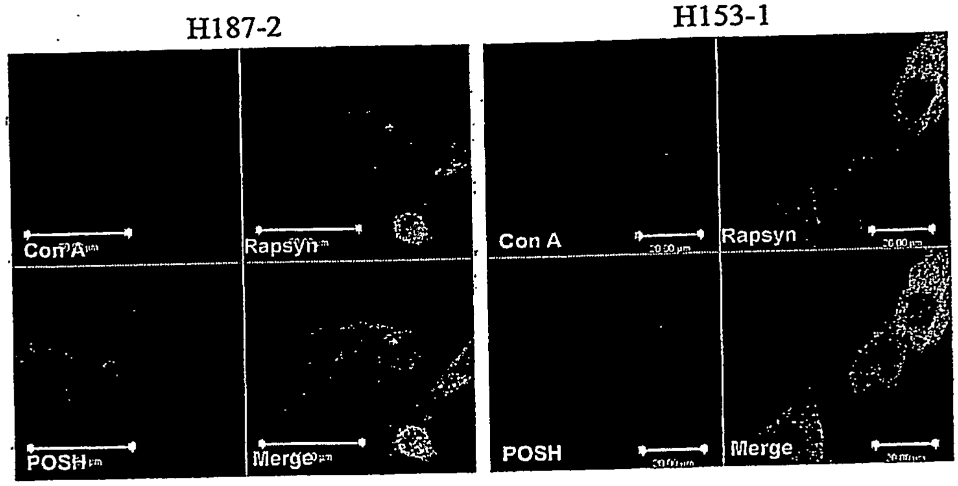


Figure 25.



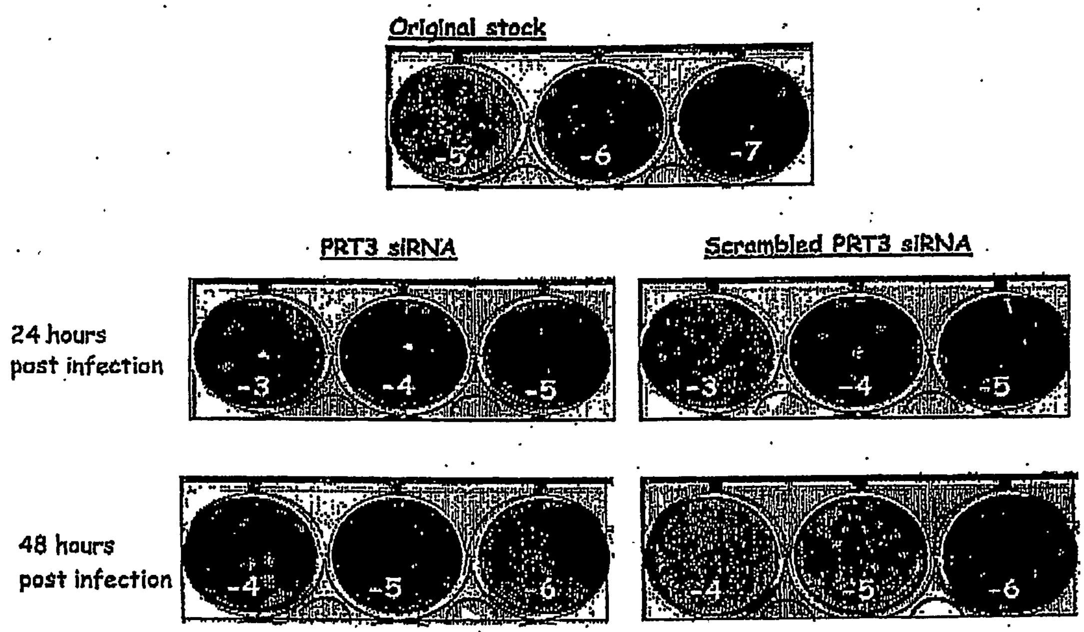
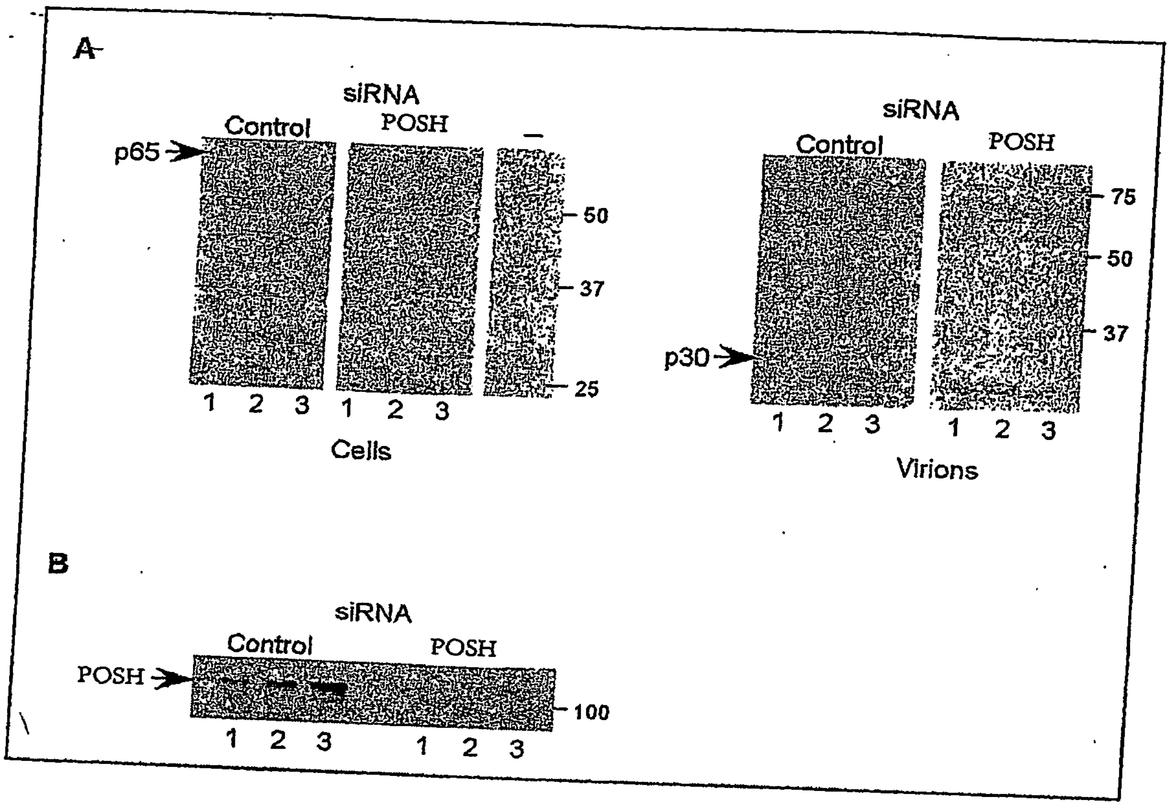

Original stock
-5
-6
-7
PRT3 siRNA
Scrambled PRT3 siRNA
24 hours
post infection
-3
-4
-5
-3
-4
-5
48 hours
post infection
-4
-5
-6
-4
-5
-6

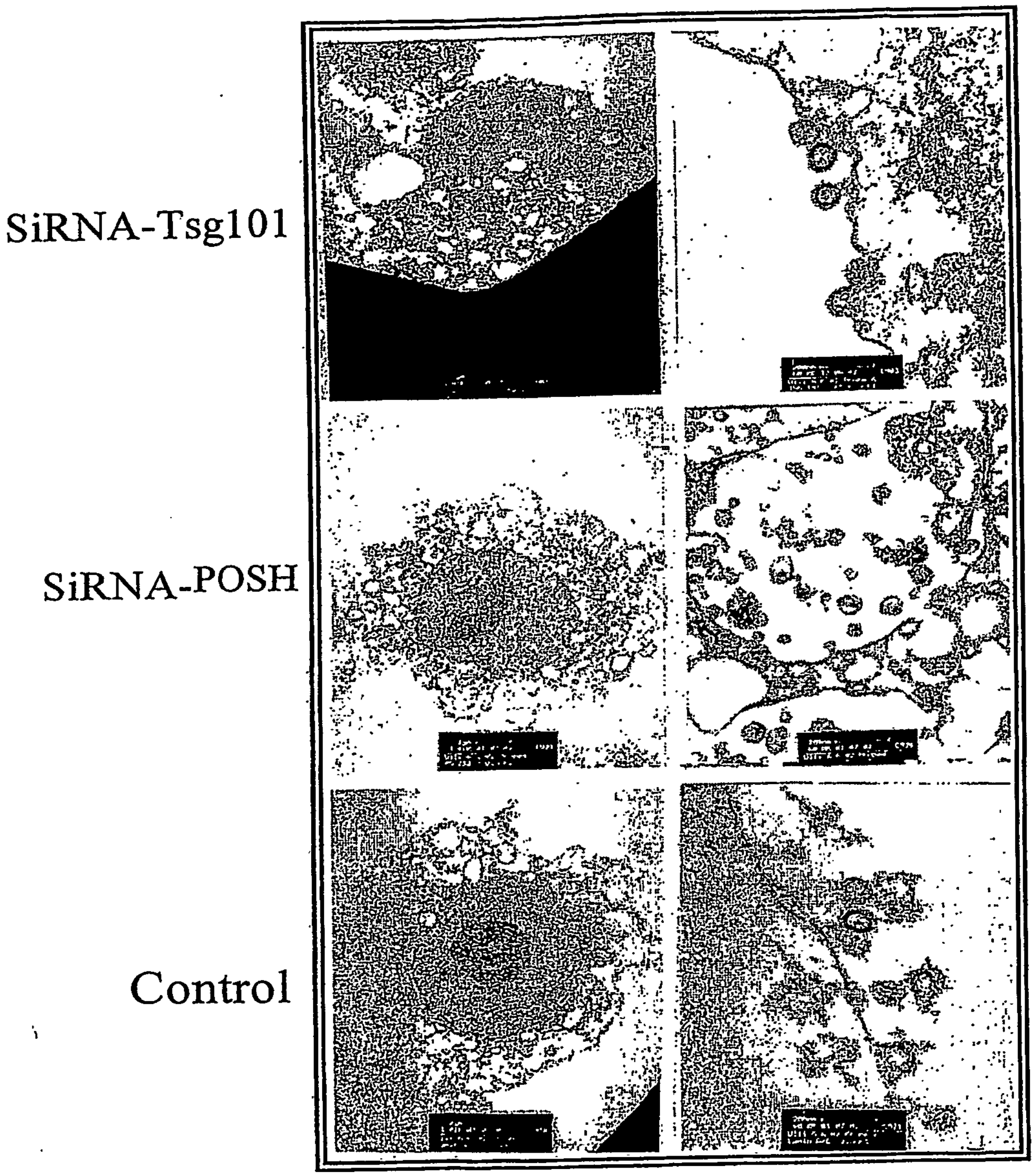
FIGURE 26

Figure 27.


A
siRNA
Control
POSH
—
p65
50
37
25
1 2 3 1 2 3
Cells
siRNA
Control
POSH
75
50
37
p30
1 2 3 1 2 3
Virions
B
siRNA
Control
POSH
POSH
100
1 2 3 1 2 3

PCT/US04/06308

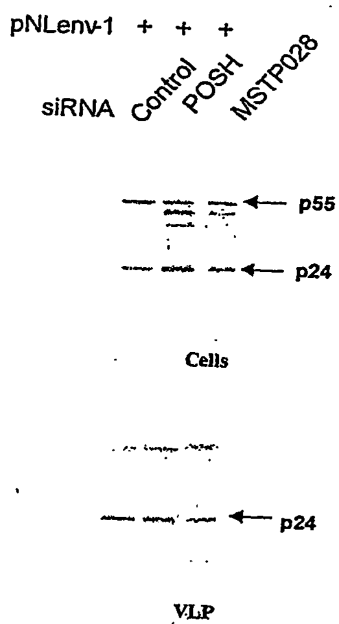
Figure 28.



5

PCT/US04/06308

Figure 29A.



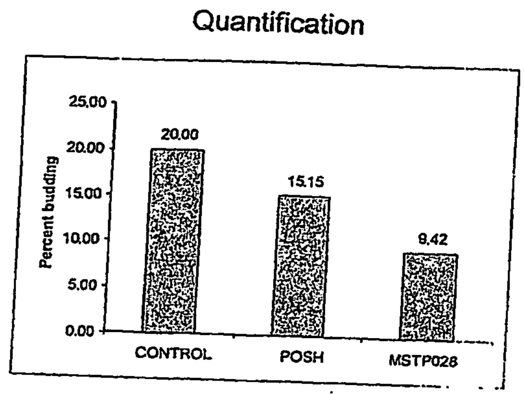
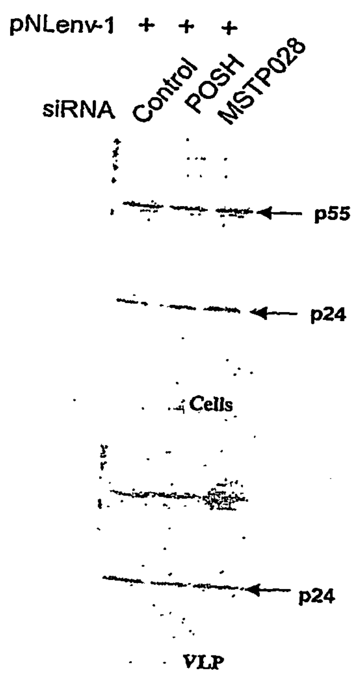


Figure 29B.


pNLenv-1 + + +
siRNA Control POSH MSTP028
p55
p24
Cells
p24
VLP

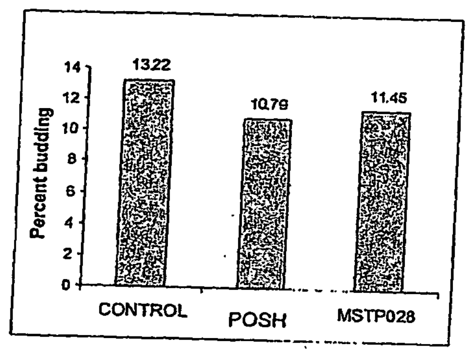

Percent budding
14
12
10
8
6
4
2
0
13.22
10.79
11.45
CONTROL
POSH
MSTP028

Figure 30. Putative PKA phosphorylation sites in hPOSH.

MDESALLDLLECPVCLERLDASAKVLPCQHTFCKRCLLGIVGSRNELRCPECRTLVGSGVEELPSNILLV
RLLDGIKQRPWKPGPGGGSGTNCTNALRSQSSTVANCSSKDLQSSQGGQQPRVQSWSPPVRGIPQLPCAK
ALYNYEGKEPGDLKFSKGDIIILRRQVDENWYHGEVNGIHGFFPTNFVQIIKPLPQPPPQCKALYDFEVK
DKEADKDCLPFAKDDVLTVIRRVDENWAEGMLADKIGIFPISYVEFNSAAKQLIEWDKPPVPGVDAGECS
SAAAQSSTAPKHSDTKKNSKKRHSFTSLTMANKSSQASQNRHSMEISPPVLISSSNPTAAARISELSGL
S
CSAPSQVHISTTGLIVTPPPSSPVTTGPSFTFPSDVPYQAALGTLNPPLPPPPLLAATVLASTPPGATAA
AAAAGMGPRPMAGSTDQIAHLRPQTRPSVYVAIYPYTPRKEDELELRKGEMFLVFERCQDGWFKGTSMHT
SKIGVFPGNYVAPVTRAVTNASQAKVPMSTAGQTSRGVTMVSPSTAGGPAQKLQGNGVAGSPSVVPAAVV
SAAHIQTSPQAKVLLHMTGQMTVNQARNAVRTVAAHNQERPTAAVTPIQVQNAAGLSPASVGLSHHSLAS
PQPAPLMPGSATHTAAISISRASAPLACAAAAPLTSPSITSASLEAEPSGRIVTVLPGLPTSPDSASSAC
GNSSATKPDKDSKKEKKGLLKLLSGASTKRKPRVSPPASPTLEVELGSAELPLQGAVGPELPPGGGHGRA
GSCPVDGDGPVTTAVAGAALAQDAFHRKASSLDSAVPIAPPPRQACSSLGPVLNESRPVVCERHRVVVSY
PPQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKTGLFPGSFVENI

Figure 31. Phosphorylation of hPOSH regulates binding of GTP-loaded Rac-1.

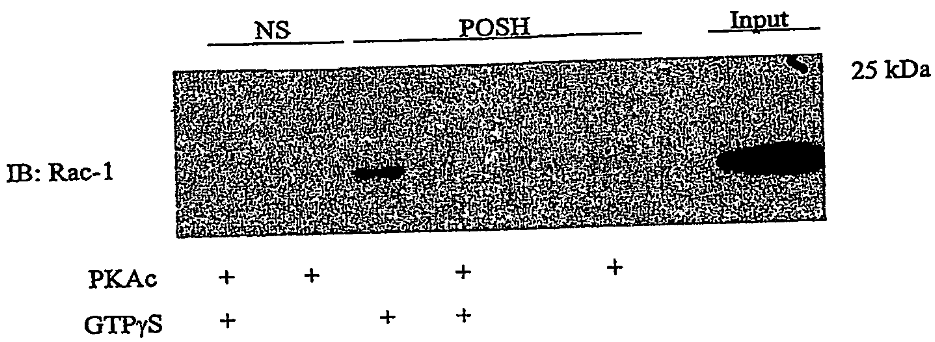
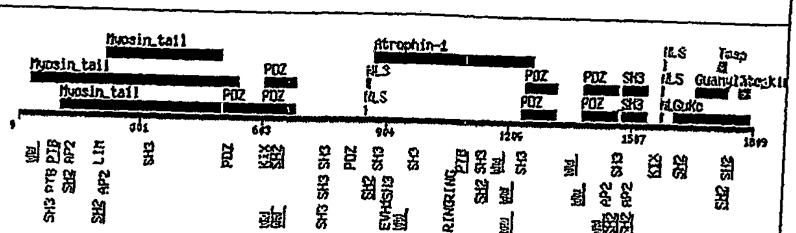
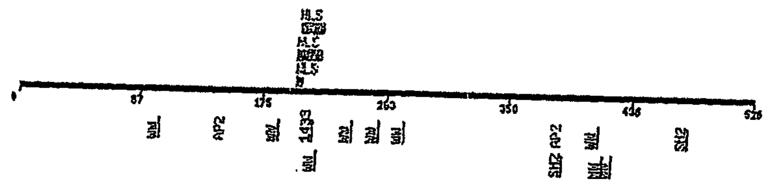
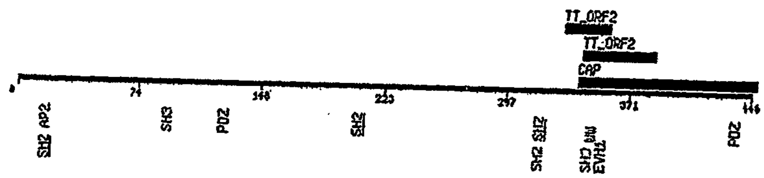
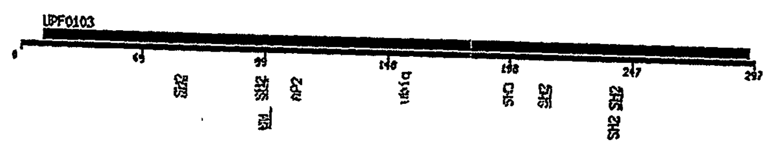
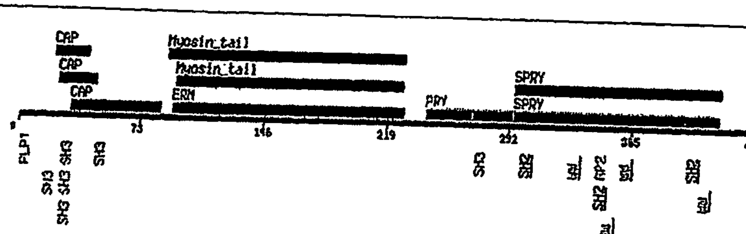
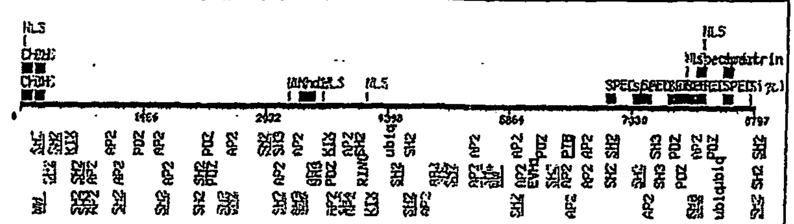
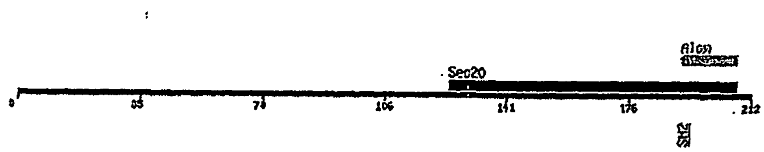
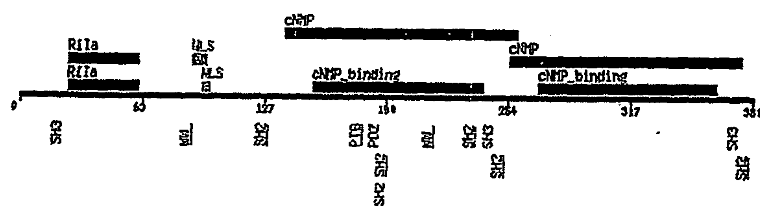
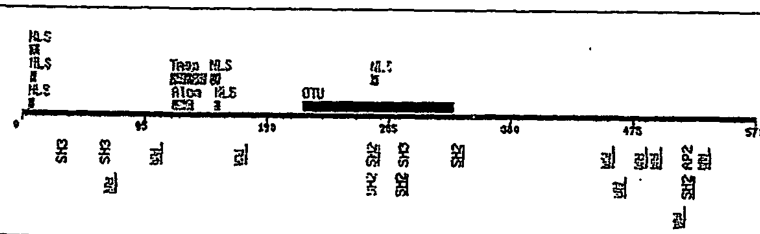
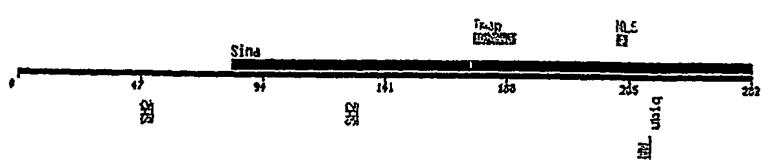
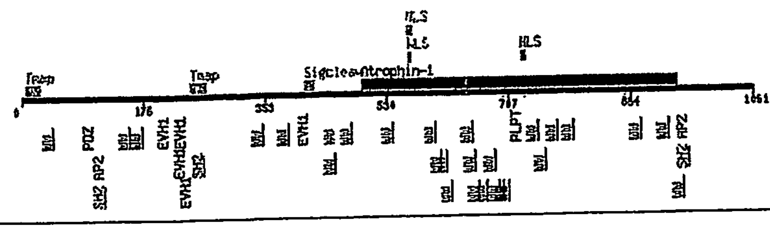
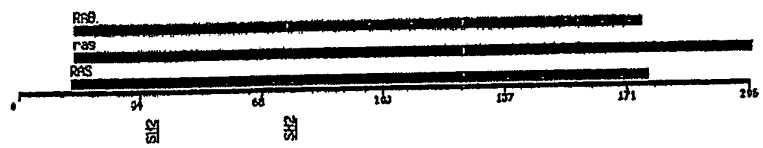
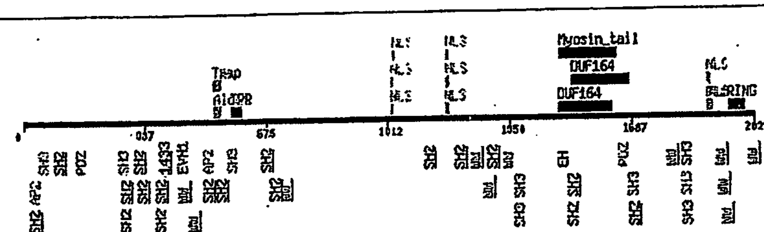
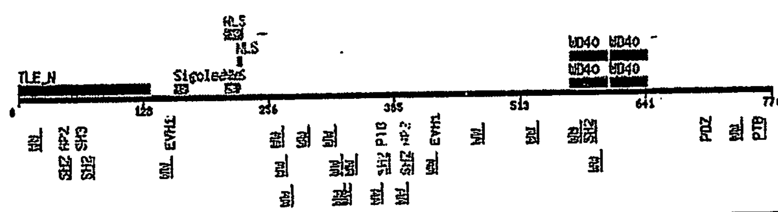
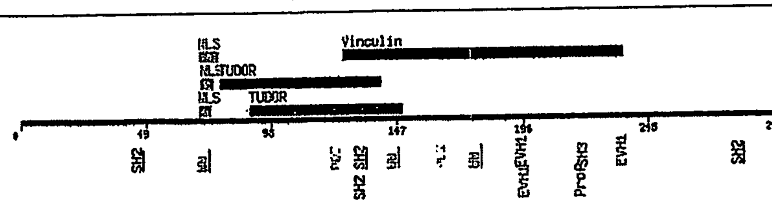
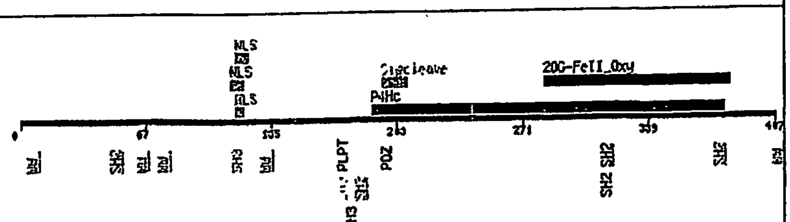


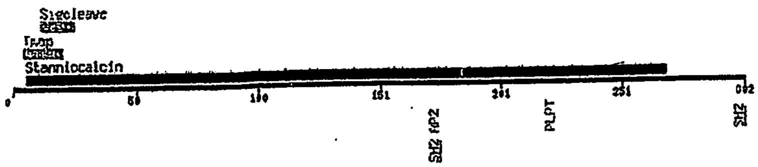
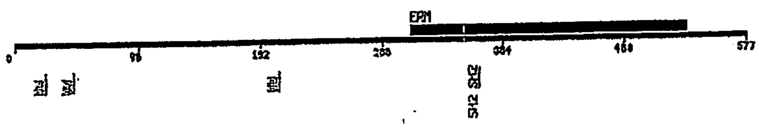
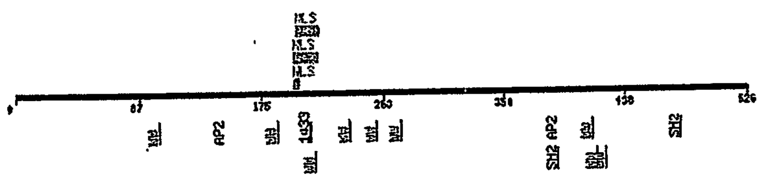
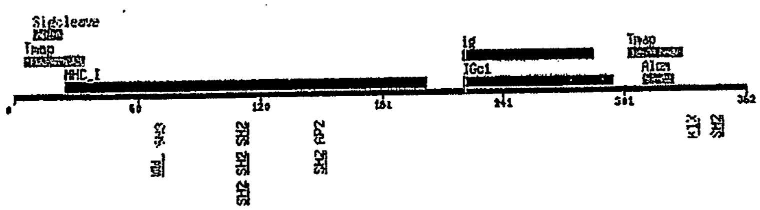
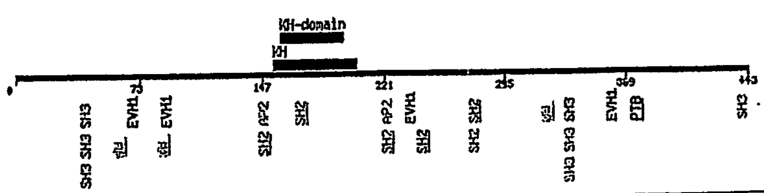
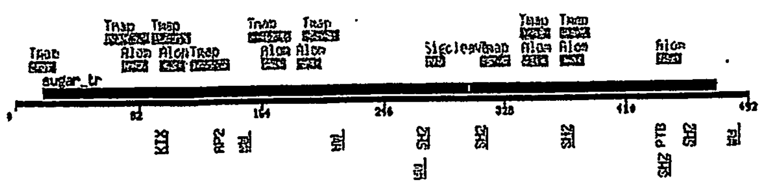
Figure 32.

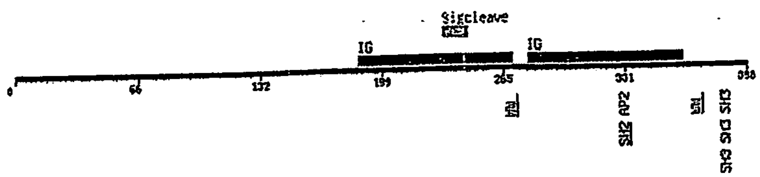
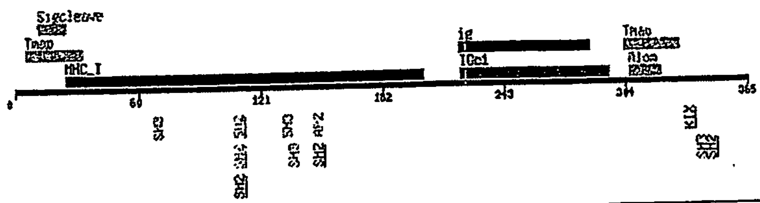
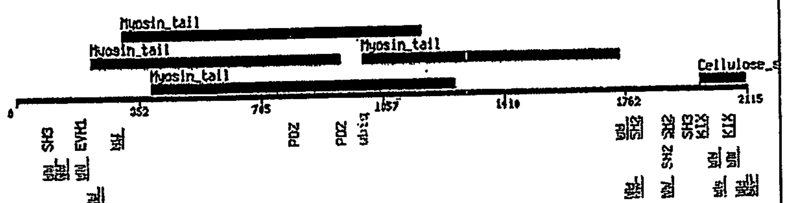
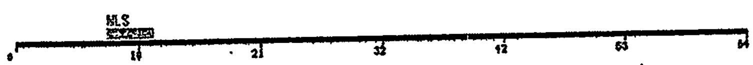
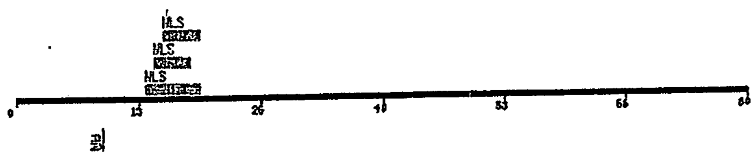
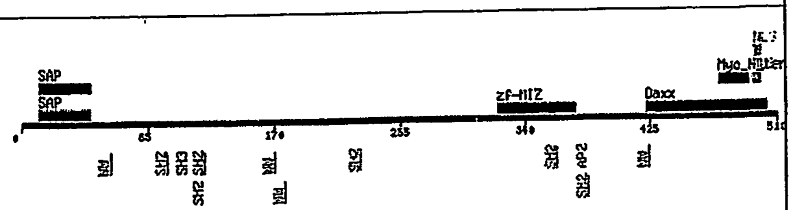
BLAST hit	UniGene	Name	Longest Protein	Domain Analysis
AK092170	Hs.302746	**MSTP028**		
AB011155.1	Hs.170290	**DLG5** discs, large (Drosophila) homolog 5	NP_004738 aa887	
XM_208944.1	None		XP_208944.1	
AB046818	Hs.23740	**KIAA1598** KIAA1598 protein	10047271 aa146	
BC018733.1	Hs.20814	**CGI-27** C21orf19-like protein	4680693	
AL080170.1	Hs.51692	**BIA2** BIA2	5262640	
BC036531.1	Hs.172928	**COL1A1** collagen, type I, alpha 1		
J03930.1		Human intestinal alkaline phosphat		

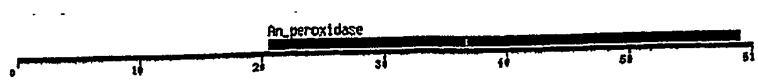
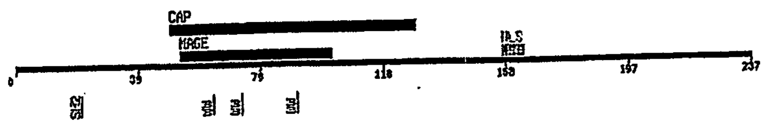
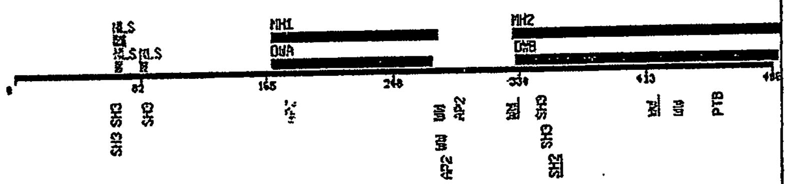
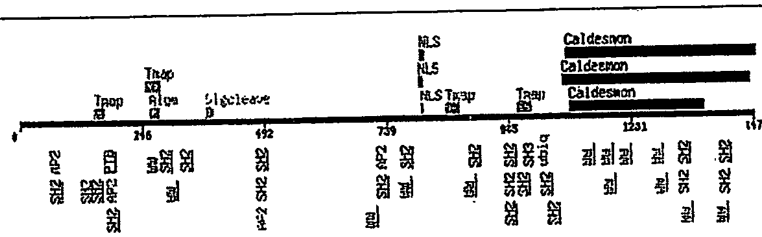
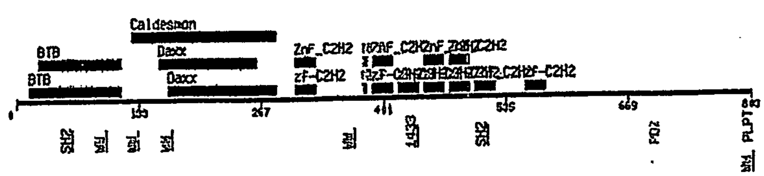
BLAST hit	UniGene	Name	Longest Protein	Domain Analysis
AF535142 AF535142	Hs.416719	SYNE1 spectrin repeat containing, nuclear envelope 1	AAN6044 2.1 8797 aa	
M93425	Hs.62	PTPN12 protein tyrosine phosphatase, non-receptor type 12	292409 aa504>	
BC009710	Hs.100651	GOSR2 golgi SNAP receptor complex member 2	16905522 16905520	
M18468 M18468 BC036285 M18468	Hs.183037	PRKAR1A protein kinase, cAMP-dependent, regulatory, type I, alpha (tissue specific extinguisher 1)		
AL137509 in 3'UTR?	Hs.184029	DKFZp761A052 hypothetical protein	AAH09917	
BC013082 U76247	Hs.295923	SIAH1 seven in absentia homolog 1 (Drosophila)	AAC51907	
BC032851	Hs.3144	CBLB Cas-Br-M (murine) ecotropic		

PCT/US04/06308

BLAST hit	UniGene	Name	Longest Protein	Domain Analysis
		retroviral transforming sequence b		
BC006358 -bp 2026 bp 1561 bp1564 bp1562 bp1561 bp1564	Hs.66048	**VCY2IP1** VCY2 interacting protein 1	21739763	
BC039858	Hs.6906	**RALA** v-ral simian leukemia viral oncogene homolog A (ras related)	24980847 aa1>	
D83077	Hs.118174	**TTC3** tetratricopeptide repeat domain 3	1304132 aa1027 aa1030	
M99435	Hs.28935	**TLE1** transducin-like enhancer of split 1 (E(sp1) homolog, Drosophila)	307510	
U18423	Hs.288986	**SMN1** survival of motor neuron 1, telomeric	624186	
BC001723, AJ310544	Hs.324277	**EGLN2** egl nine homolog 2 (C. elegans)	14547148	
BC000386	Hs.58189	**EIF3S3** eukaryotic translation		

BLAST hit	UniGene	Name	Longest Protein	Domain Analysis
		initiation factor 3, subunit 3 gamma, 40kDa		
AF055460	Hs.155223	**STC2** stanniocalcin 2	AAC27036	
BC013876	Hs.278898	**OPTN** optineurin	AAH13876	
XM_208944 AK094466	Hs.420088	Unnamed protein product	XP_208944	
X61709	Hs.77961	**HLA-B** major histocompatibility complex, class I, B	32189	
M88108	Hs.119537	**KHDRBS1** KH domain containing, RNA binding, signal transduction associated 1	189500	
K03195/ NM_006516	Hs.169902	**SLC2A1** solute carrier family 2 (facilitated glucose transporter),	5730051	

BLAST hit	UniGene	Name	Longest Protein	Domain Analysis
AL137493	Hs.35945	**DKFZp434B1231** hypothetical protein DKFZp434B1231	6808117	
L06425	Hs.181244	**HLA-A**	575249	
BC008345	Hs.301512	**NUMA1** nuclear mitotic apparatus protein 1	14249928 963aa 35119 2115aa	
AF077202 AF077202	Hs.397853	**HSPC016** hypothetical protein HSPC016	12654537 64aa	
BC000449	Hs.183704	**UBC**		
D26121	Hs.169303	ZFM1 protein alternatively spliced product domain A, B and G		
AF077952	Hs.105779	**PIASY** protein inhibitor of activated STAT protein PIASy	3643111	

BLAST hit	UniGene	Name	Longest Protein	Domain Analysis
BC007034	Hs.118786	**MT2A** metallothionein 2A	13937857	
AF293026	Hs.32587	**SRA1** steroid receptor RNA activator 1	9930614	
X66899	Hs.129953	**EWSR1** Ewing sarcoma breakpoint region 1		Synaptophysinx4; Transcription factor IIA; zinc finger x4; NLSx3,
AF035528	Hs.153863	**MADH6** MAD, mothers against decapentaplegic homolog 6 (Drosophila)	2736316	
AF441770	Hs.16411	**THOC2** THO complex 2	AAM28436	
Y09723	Hs.33532	**ZNF151** zinc finger protein 151 (pHZ-67)	2230871	

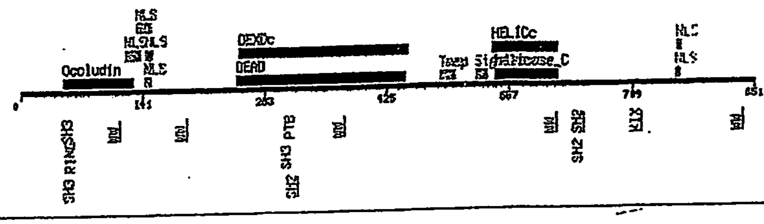
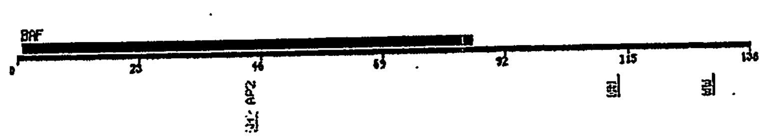
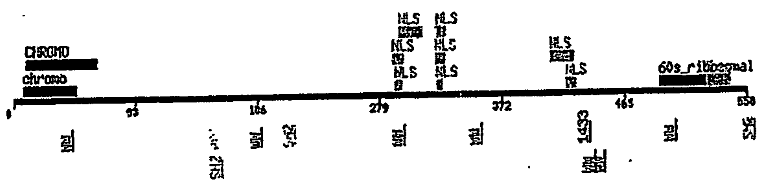
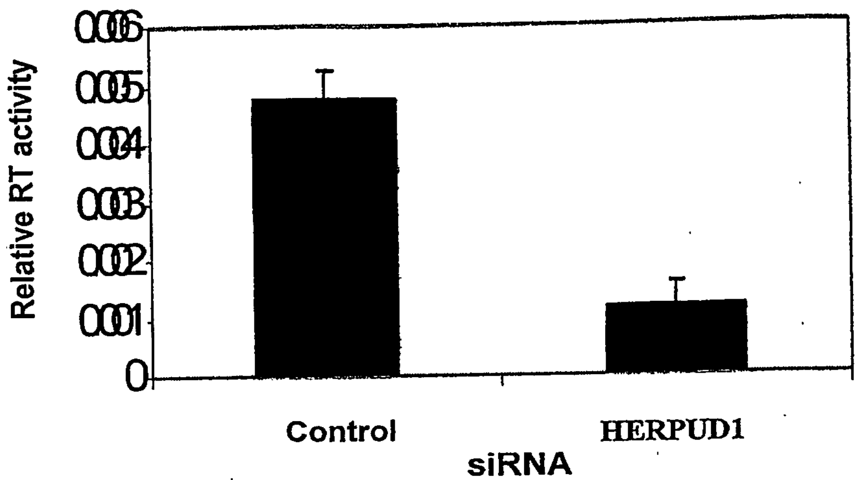
BLAST hit	UniGene.	Name	Longest Protein	Domain Analysis
BC012726	Hs.69331	**DDX31** DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 31	7505907	
NM_032958	Hs.375569	**POL R2J2** DNA directed RNA polymerase II polypeptide J-related gene		
AF068235.1	Hs.433759	**BANF1** barrier to autointegration factor 1	3002951	
BC014967.1	Hs.5637	**CBX4** chromobox homolog 4	4502603 aa319	

Figure 33.



Relative RT activity
0.006
0.005
0.004
0.003
0.002
0.001
0
Control
HERPUD1
siRNA

Figure 34A.

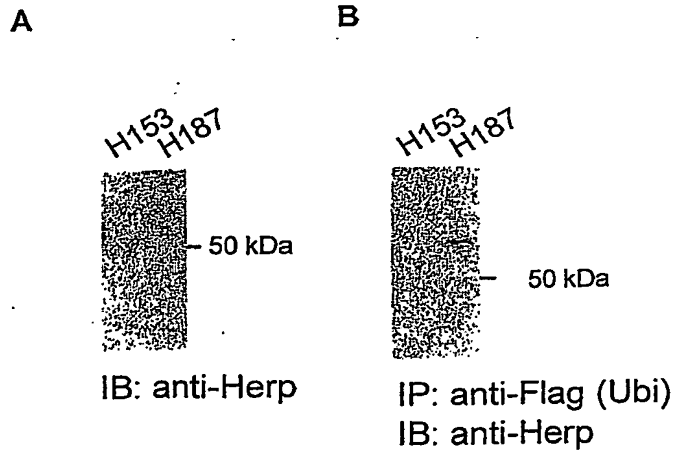
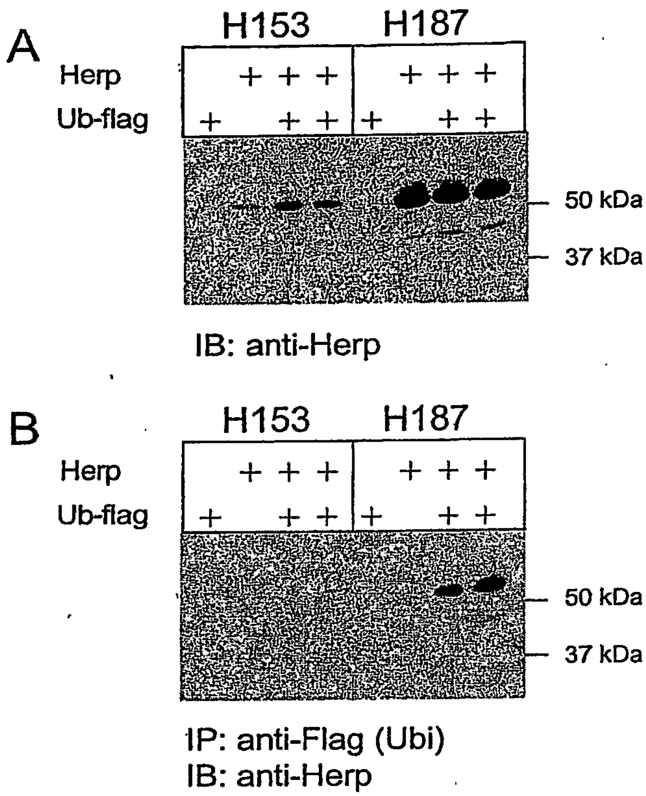

A
H153
H187
50 kDa
IB: anti-Herp
B
H153
H187
50 kDa
IP: anti-Flag (Ubi)
IB: anti-Herp

Figure 34B.


A
H153
H187
Herp
Ub-flag
50 kDa
37 kDa
IB: anti-Herp
B
H153
H187
Herp
Ub-flag
50 kDa
37 kDa
IP: anti-Flag (Ubi)
IB: anti-Herp

PCT/US04/06308

Figure 35.

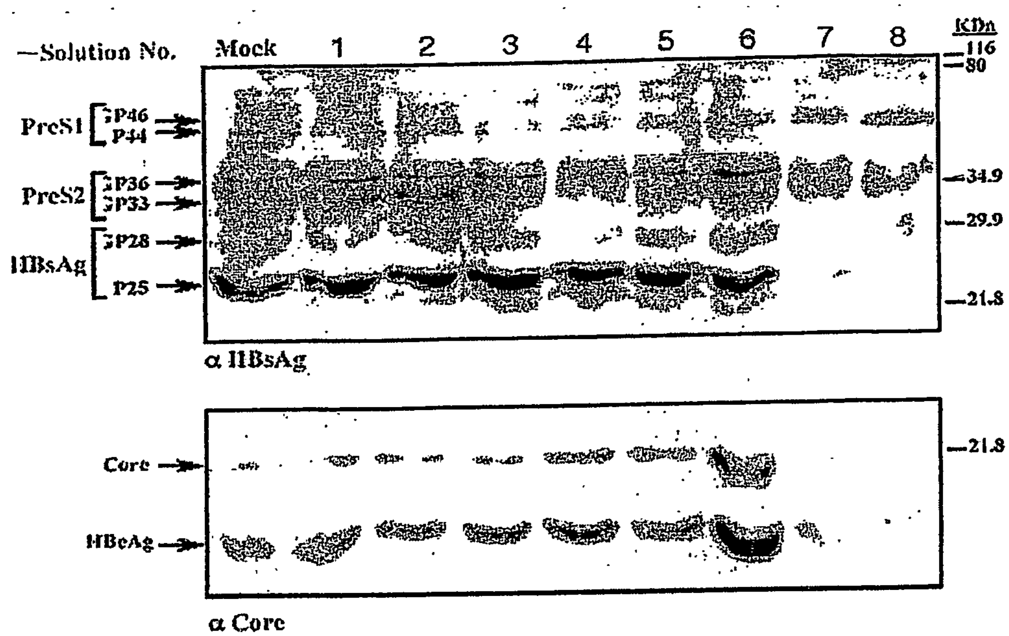


FIGURE 36

Unigene Name: Arf1 **Unigene ID:** Hs.286221

Human Arf1 mRNA sequence - var1 (public gi: 3360490) (SEQ ID NO: 325)
GCAAAACCAACGCCTGGCTCGGAGCAGCAGCCTCTGAGGTGTCCCTGGCCAGTGTCCTTCCACCTGTCCA
CAAGCATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCATCCTCAT
GGTGGGCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTGACCACC
ATTCCCACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGGACGTGG
GTGGCCAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTTCGTGGT
GGACAGCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAGGACGAG
CTCCGGGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGGCCGAGA
TCACAGACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGCGCCACCAG
CGGCGACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTGAACGCGACCC
CCCTCCCTCTCACTCCTCTTGCCCTCTGCTTTACTCTCATGTGGCAAACGTGCGGCTCGTGGTGTGAGTG
CCAGAAGCTGCCTCCGTGGTTTGGTCACCGTGTGCATCGCACCGTGCTGTAAATGTGGCAGACGCAGCCT
GCGGCCAGGCTTTTTATTTAATGTAAATAGTTTTTGTTTCCAATGAGGCAGTTTCTGGTACTCCTATGCA
ATATTACTCAGCTTTTTTTATTGTAAAAAGAAAAATCAACTCACTGTTCAGTGCTGAGAGGGGATGTAGG
CCCATGGGCACCTGGCCTCCAGGAGTCGCTGTGTTGGGAGAGCCGGCCACGCCCTTGGCTTTAGAGCTGT
GTTGAAATCCATTTTGGTGGTTGGTTTTTAACCCAAACTCAGTGCATTTTTTAAAATAGTTAAGAATCCA
AGTCGAGAACACTTGAACACACAGAAGGGAGACCCCGCCTAGCATAGATTTGCAGTTACGGCCTGGATGC
CAGTCGCCAGCCCAGCTGTTCCCCTCGGGAACATGAGGTGGTGGTGGCGCAGCAGACTGCGATCAATTCT
GCATGGTCACAGTAGAGATCCCCGCAACTCGCTTGTCCTTGGGTCACCCTGCATTCCATAGCCATGTGCT
TGTCCCTGTGCTCCCACGGTTCCCAGGGGCCAGGCTGGGAGCCCACAGCCACCCCACTATGCCGCAGGCC
GCCCTACCCACCTTCAGGCAGCCTATGGGACGCAGGGCCCCATCTGTCCCTCGGTCGCCGTGTGGCCAGA
GTGGGTCCGTCGTCCCCAACACTCGTGCTCGCTCAGACACTTTGGCAGGATGTCTGGGGCCTCACCAGCA
GGAGCGCGTGCAAGCCGGGCAGGCGGTCCACCTAGACCCACAGCCCCTCGGGAGCACCCCACCTCTGTGT
GTGATGTAGCTTTCTCTCCCTCAGCCTGCAAGGGTCCGATTTGCCATCGAAAAAGACAACCTCTACTTTT
TTCTTTTGTATTTTGATAAACACTGAAGCTGGAGCTGTTAAATTTATCTTGGGGAAACCTCAGAACTGGT
CTATTTGGTGTCGTGGAACCTCTTACTGCTTTCAATACACGATTAGTAATCAAAAAAAAAAAAAAAAAAA
AAAAAAA

Human Arf1 mRNA sequence - var2 (public gi: 30583624) (SEQ ID NO: 326)
ATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCATCCTCATGGTGG
GCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTGACCACCATTCC
CACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGGACGTGGGTGGC
CAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTTCGTGGTGGACA
GCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAGGACGAGCTCCG
GGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGGCCGAGATCACA
GACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGCGCCACCAGCGGCG
ACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTAG

Human Arf1 mRNA sequence - var3 (public gi: 34527605) (SEQ ID NO: 327)
AAAACCAACGCCTGGCTCGGAGCAGCAGCCTCTGAGGTGTCCCTGGCCAGTGTCCTTCCACCTGTCCACA
AGCATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCATCCTCATGG
TGGGCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTGACCACCAT
TCCCACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGGACGTGGGT
GGCCAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTTCGTGGTGG
ACAGCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAGGACGAGCT
CCGGGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGGCCGAGATC
ACAGACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGTGCCACCAGCG
GCGACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTGAACGCGACCCCC
CTCCCTCTCACTCCTCTTGCCCTCTGCTTTACTCTCATGTGGCAAACGTGCGGCTCGTGGTGTGAGTGCC
AGAAGCTGCCTCCGTGGTTTGGTCACCGTGTGCATCGCACCGTGCTGTAAATGTGGCAGACGCAGCCTGC
GGCCAGGCTTTTTATTTAATGTAAATAGTTTTTGTTTCCAATGAGGCAGTTTCTGGTACTCCTATGCAAT
ATTACTCAGCTTTTTTTATTGTAAAAAGAAAAATCAACTCACTGTTCAGTGCTGAGAGGGGATGTAGGCCC
ATGGGCACCTGGCCTCCAGGAGTCGCTGTGTTGGGAGAGCCGGCCACGCCCTTGGCTTTAGAGCTGTGTT
GAAATCCATTTTGGTGGTTGGTTTTTAACCCAAACTCAGTGCATTTTTTAAAATAGTTAAGAATCCAAGT
CGAGAACACTTGAACACACAGAAGGGAGACCCCGCCTAGCATAGATTTGCAGTTACGGCCTGGATGCCAG
TCGCCAGCCCAGCTGTTCCCCTCGGGAACATGAGGTGGTGGTGGCGCAGCAGACTGCGATCAATTCTGCA
TGGTCACAGTAGAGATCCCCGCAACTCGCTTGTCCTTGGGTCACCCTGCATTCCATAGCCATGTGCTTGT

CCCTGTGCTCCCACGGTTCCCAGGGGCCAGGCTGGGAGCCCACAGCCACCCCACTATGCCGCAGGCCGCC
CTACCCACCTTCAGGCAGCCTATGGGACGCAGGGCCCCATCTGTCCCTCGGTCGCCGTGTGGCCAGAGTG
GGTCCGTCGTCCCCAACACTCGTGCTCGCTCAGACACTTCGGCAGGATGTCTGGGGCCTCACCAGCAGGA
GCGCGTGCAAGCCGGGCAGGCGGTCCACCTAGACCCACAGCCCCTCGGGAGCACCCCACCTCTGTGTGTG
ATGTAGCTTTCTCTCCCTCAGCCTGCAAGGGTCCGATTTGCCATCGAAAAAGACAACCTCTACTTTTTTC
TTTTGTATTTTGATAAACACTGAAGCTGGAGCTGTTAAATTTATCTTGGGGAAACCTCAGAACTGGTCTA
TTTGGTGTCGTGGAACCTCTTACTGCTTTCAATACACGATTAGTAATCAACTGTTTTGTATACTTGTTTT
CAGTTTTCATTTCGACAAACAAGCACTGTAATTATAGCTATTAGAATAAAATCTCTTAACTATTT

Human Arf1 mRNA sequence - var4 (public gi: 6995997) (SEQ ID NO: 328)
GCAAAACCAACGCCTGGCTCGGAGCAGCAGCCTCTGAGGTGTCCCTGGCCAGTGTCCTTCCACCTGTCCA
CAAGCATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCATCCTCAT
GGTGGGCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTGACCACC
ATTCCCACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGGACGTGG
GTGGCCAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTTCGTGGT
GGACAGCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAGGACGAG
CTCCGGGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGGCCGAGA
TCACAGACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGCGCCACCAG
CGGCGACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTGAACGCGACCC
CCCTCCCTCTCACTCCTCTTGCCCTCTGCTTTACTCTCATGTGGCAAACGTGCGGCTCGTGGTGTGAGTG
CCAGAAGCTGCCTCCGTGGTTTGGTCACCGTGTGCATCGCACCGTGCTGTAAATGTGGCAGACGCAGCCT
GCGGCCAGGCTTTTTATTTAATGTAAATAGTTTTTGTTTCCAATGAGGCAGTTTCTGGTACTCCTATGCA
ATATTACTCAGCTTTTTTTATTGTAAAAAGAAAAATCAACTCACTGTTCAGTGCTGAGAGGGGATGTAGG
CCCATGGGCACCTGGCCTCCAGGAGTCGCTGTGTTGGGAGAGCCGGCCACGCCCTTGGCTTAGAGCTGTG
TTGAAATCCATTTTGGTGGTTGGTTTTAACCCAAACTCAGTGCATTTTTTAAAATAGTTAAGAATCCAAG
TCGAGAACACTTGAACACACAGAAGGGAGACCCCGCCTAGCATAGATTTGCAGTTACGGCCTGGATGCCA
GTCGCCAGCCCAGCTGTTCCCCTCGGGAACATGAGGTGGTGGTGGCGCAGCAGACTGCGATCAATTCTGC
ATGGTCACAGTAGAGATCCCCGCAACTCGCTTGTCCTTGGGTCACCCTGCATTCCATAGCCATGTGCTTG
TCCCTGTGCTCCCACGGTTCCCAGGGGCCAGGCTGGGAGCCCACAGCCACCCCACTATGCCGCAGGCCGC
CCTACCCACCTTCAGGCAGCCTATGGGACGCAGGCCCCATCTGTCCCTCGGTCCGCGTGTGGCCAGAGTG
GTCCGTCGTCCCCAACACTCGTGCTCGCTCAGACACTTTGGCAGGATGTCTGGGGCCTCACCAGCAGGAG
CGCGTGCAAGCCGGGCAGGCGGTCCACCTAGACCCACAGCCCCTCGGGAGCACCCCACCTCTGTGTGTGA
TGTAGCTTTCTCTCCCTCAGCCTGCAAGGGTCCGATTTGCCATCGAAAAAGACAACCTCTACTTTTTTCT
TTTGTATTTTGATAAACACTGAAGCTGGAGCTGTTAAATTTATCTTGGGGAAACCTCAGAACTGGTCTAT
TTGGTGTCGTAGGAACCTCTTACTGCTTTCAATACACGATTAGTAATCAACTGTTTTGTATACTTGTTTT
CAGTTTTCATTTCGACAAACAAGCACTGTAATTATAGCTATTAGAATAAAATCTCTTAACTATT

Human Arf1 mRNA sequence - var5 (public gi: 7020834) (SEQ ID NO: 329)
CCTTACCCGGCGTGCCCCGCGCCCGGAGGCGCTGACGTGGCCGCCGTCAGAGCCGCCATCTTGTGGGAGC
AAAACCAACGCCTGGCTCGGAGCAGCAGCCTCTGAGGTGTCCCTGGCCAGTGTCCTTCCACCTGTCCACA
AGCATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCATCCTCATGG
TGGGCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTGACCACCAT
TCCCACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGGACGTGGGT
GGCCAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTTCGTGGTGG
ACAGCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAGGACGAGCT
CCGGGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGGCCGAGATC
ACAGACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGCGCCACCAGCG
GCGACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTGAACGCGACCCCC
CTCCCTCTCACTCCTCTTGCCCTCTGCTTTACTCTCATGTGGCAAACGTGCGGCTCGTGGTGTGAGTGCC
AGAAGCTGCCTCCGTGGTTTGGTCACCGTGTGCATCGCACCGTGCTGTAAATGTGGCAGACGCAGCCTGC
GGCCAGGCTTTTTATTTAATGTAAATAGTTTTTGTTTCCAATGAGGCAGTTTCTGGTACTCCTATGCAAT
ATTACTCAGCTTTTTTTATTGTAAAAAGAAAAATCAACTCACTGTTCAGTGCTGAGAGGGGATGTAGGCC
CATGGGCACCTGGCCTCCAGGAGTCGCTGTGTTGGGAGAGCCGGCCACGCCCTTGGCTTTAGAGCTGTGT
TGAAATCCATTTTGGTGGTTGGTTTTTAACCCAAACTCAGTGCATTTTTTAAAATAGTTAAGAATCCAAG
TCGAGAACACTTGAACACACAGAAGGGAGACCCCGCCTAGCATAGATTTGCAGTTACGGCCTGGATGCCA
GTCGCCAGCCCAGCTGTTCCCCTCGGGAACATGAGGTGGTGGTGGCGCAGCAGACTGCGATCAATTCTGC
ATGGTCACAGTAGAGATCCCCGCAACTCGCTTGTCCTTGGGTCACCCTGCATTCCATAGCCATGTGCTTG
TCCCTGTGCTCCCACGGTTCCCAGGGGCCAGGCTGGGAGCCCACAGCCACCCCACTATGCCGCAGGCCGC
CCTACCCACCTTCAGGCAGCCTATGGGACGCAGGGCCCCATCTGTCCCTCGGTCGCCGTGTGGCCAGAGT
GGGTCCGTCGTCCCCAACACTCGTGCTCGCTCAGACACTTTGGCAGGATGTCTGGGGCCTCACCAGCAGG
AGCGCGTGCAAGCCGGGCAGGCGGTCCACCTAGACCCACAGCCCCTCGGGAGCACCCCACCTCTGTGTGT
GATGTAGCTTTCTCTCCCTCAGCCTGCAAGGGTCCGATTTGCCATCGAAAAAGACAACCTCTACTTTTTT

CTTTTGTATTTTGATAAACACTGAAGCTGGAGCTGTTAAATTTATCTTGGGGAAACCTCAGAACTGGTCT
ATTTGGTGTCGTGGAACCTCTTACTGCTTTCAATACACGATTAGTAATCAACTGTTTTGTATACTTGTTT
TCAGTTTTCATTTCGACAAACAAGCACTGTAATTATAGCTATTAGAATAAAATCTCTTAACTATTAAAAA
AAAAAAAAAAAAAAAAAAAAA

Human Arf1 mRNA sequence - var6 (public gi: 10435849) (SEQ ID NO: 330)

AGCTCAGTGCCCAGCATGTCTGTGGTGAGTGTGTAGTTCAGGAAGTGAACTGGCAAAACTGAGTATCACC
CTCTCTTCCTGGGTTCTTGCCACTCCCCTGAAAACCAGGGTAGCATTGTCACATCAGATAGCTCCGCTAC
GTGTGCGCTGACCATGCTGAGATGGGCACTGTGGACTCAGCCTCTGGTCATTGCTGGAACCAGCGGCCTC
CATGTGAGGTACAGGGGAACGCACTGCTAGCAGATGGTTGGGATGTGGACACTCGTCCTGCCCTCTTGGC
TTGGTGCTGTGCCATCGCACAGTCATTCGCTGTTTAGCATGCATGGGAGAGAGTGAAGCACAAGGGCCCA
GGCCCCTGGGAGTGCCTGCCCTCAATTTGGAAGAGCCCTTGGGCACAGCATAGGCGCCTGGCAGAATTGG
ACTGGGCCATGATCCAGGGCATTGGGACCTCACCTAGGAGTTGGGGTTCTGGTCAGAAGCCCTGTGGAGA
CAGGGTCTCCCCTGTGGGCACCAAACTGACCTCAAACTGCTGGTTCTTTGGCCCTGGGGACAGGGCTGGT
TGAAGTACTCTCCCGGCAGCTGTCACCTGCAGGGAGAGGTGGGGGTAGGGGTGCTGTGTTTCTTAGCTGT
TCCTCGTTGCAGTGTTAAATCCCTGCAGGTTCCTTATTCTCAGCTTGTGTTGTGAGTTTCAGTGTTGGGG
GCTAATGTGGGTTTGCGTTTTTGGTCTTGGTTTTCCCAGTGGCCAGTCCATCAGCCACTGCACTGGGGGC
CAGGTAGAGGCCAACTGCACCCTGCCTGCCAGAGTAGAAATACTGGTAGGCCCCAGGCTCTGCTGCCCCT
TCCATGTCCTTGTGTAAGCATCCATGGACAAAGCTGACTCACGGGGTGTGCACAGCTGCAGGGAGGCCAG
GAAACAGGGGTTTTATTCTAGAGGGCCTTGTGCTCAGTGACAGACCAGAGTCCCATCACTGAGAGAGCAG
GGCTGGGGCAGCACAAAGGACTGGATAGCATTTGCCATGATGCCATGTGCACAGCCAGTGCAAGTCCTTC
ATTGTAGCTGTGGTCAGAGGTCATGAGACACTGCCTTCAGCAGCCCTGGGAGTCCACCTGGTGTGTGCTT
AGAGCTGTGCATCTGCAGATTTCAGAAGGACTTACGTTTGGTGAGGTGCTTTGAAGTAACACTTCACAAA
TACCAAGAAGCAAGAAATACACAAATAAGCAGGTAATGGTTCTTTGGTGTTTACATTAGCTAGTGGGCAA
CGGTTCTTTGGTGTTCACATTAGCTATAGTCCCAGAACTCAGTCCATGAGGTGGAATCACAAAAATGGAA
TTCATTTCTGGCTGTCAGTACACAAACTGATTTAAGATATCACCTTGAATTTTAAGCTGACAAACAGTGA
TCTAAACTGAATTTCACTGATTGCCCCACCTGAAAGTCAGACCTGATAGATAATGCCCTCCCTTAACTCA
AGGCCAGCAGCAGATGTGTTAGAGGGGACCCTTGTGCCTCGCAGCCCTCATCTCCTAATGGCTGTGGGGT
CACTGTGTCGAGTTGTAATGCCTAATGAGCTCCTCTAAAAACATCCTGAAACTTGTGTAAAAAAACAGCA
GACTCCCAGTGGAACTCGCCTTCAGATGCAGCCCAGAATAAGAGTTCTAGAATGTGTGTGCCATCCTTTT
GTCTCAATCTGCATGATTGCAAGTCTCTTCAACATGATTGGGTGCGTGGAGTGTCTCGGTCATGTGTCTT
CCCCTCTGAGCATGCCTTTTGATTCGCACCTGTGTCACAATTGTGCCAGCCTGTGAGATGTGTCTGCCTG
TCACCAGTATCGGCACATTTAGTTTTCCCTTTACGTGAGTTTTGGTAAAATAGTGACAAAATGTAATGCA
GTGCTCAGTCACAGAAAAATGTCAGGCCTACAGAAATGGAGCATTTGGCTGGTGGGTAGCGTGATGACCA
TAGGCTTTATTTGGCTGGTGTGGTAAACAAGCAGCAGCTTGTGCAGGTGAGAATAAATGGCCATATTGCA
TTTCATTTTAAGGACTCCCTTAAAATGAAAATCTTCGTGTGGGACATGAACACAGGCTTTCACGAAATTG
ATCATCTACACTATATGTATGACTGTTGAAAGGCTGTTGTTCCTCAGAAATTCTTAAAATGTTATGTAAT
GTACATGAGTCCCTTCAGGAAGTCATCAGCTTTGTTCAGTTTCCTCAGATTAGATAGTAAAACTGAGATT
ATGAACTATAAAGATGTGTGTAATTTATCTGTCAGTGAACTTGACTTTAATAAAAGCTTTTTGAAAAAGA
ACTCTGGGTGGGGTGCATTGGCTCACACACATAGTCCCAACTACTGTGGAGGTCAGGGCAGGAGGATCAC
TGGAGCCCAAGAGTTCAAGATCAGCCTGGGCAGGATAGCGAGACCCTGTCTATAGAAAATATTAAAAATC
AGCTAGGCATGGTGGCTTGCCCTTGCATTCCCTGCCACTTGGGAGGCTGAGGTGGGAGGTTCGCTTGAGC
CCAGGAGCTCAAGGCTGCAATGGGCTGTGATCGAACCACTGAATTCCAACCTGGGTGACAGAGTGAGGCC
CTGTCTCAAAAAGAGAACTCTCGATGTCACTGGCTTTCCATGTAAGCAGAGCACATCATGTGAGCCCCAT
TCGTGGATGTCAGTCAGCAGAACAGAATCTTGGACCTGGAGCTTGTTTGTCCTGTGCTAGAGGTTGGAGG
TGTCTCTGTCTTTCTGTTGGTTCCTGTCAGTTCAGGTCACTTAGAGATTCTGTTACATACACCAGCTCTG
ACAGGTTGGGGGAGATGATCAACCTTCCGCCTGCGCCTGTTCCCTTCCCTGACTCATGCCAAAGTATCCC
TGAGATCTGCAAGGGACCGAGGACAGTACTGGCTGGTGGTCTGGGTACAGGCCACAGAGGCATCTGGACC
CCATGTGCATCTGGACCAGTTTGGTTGGATCCATTCATGGACACAAAACGGATGTGAACTCACAGAGCTA
CATTTTCTCCCTGCCCCTGTTCAGGCACAGTGAGGTGTCGGGGAATGTAGCTGCCAGAGTTGACTGTCCC
GTTCTTTGGTGTAATGCCTGAAGGCCACCTTTACCATTGGTCTGTGGTCCTCACTGAAGAAAGAAACATT
CTTCCTAAAAGACTTTTTTTCCTCAGAGTTGGAGCCCACAGCGTGGTCAGGAAAGAGAAGTAGCCACTGG
TGGCTCCTGGCATCCTCCTGCTGGGCAGCCCCTTCTCAAAGTGTGAGGGGTCCCCTTGTGTAGAAGCAGG
AAGGCTCTGAGAAAGTCAGGTTTGCTCCTACCACAGGATAATTCCGATGAACCTGAAAAGCGGGTTTTGG
CTTGTGTGCAGGGACTCTGGTGGAAGAAAGGGTGACAGCACCTGGCCTGGGCATGACACAAGTTAGGACC
CGTACCAAGAGGCCCTGGAATTGAGGGTGGGGGTTGCTGTGGACTCTTTCTCCCTCTTAGGAAACTCTAT
TGGGTCTCCATCTGTCACAGAAGCAGTAAATGATGTAGGGGCTGCCAGGTATAGGGTCCTGTGGGGATGC
TGGAACATGCCGAGGCAGGACGTGCCAGCCACCCTCTGCCCATATGTGCAGCAGGGCCACAGATGTGCTT
GTCGGTAGGAGAGACCAAGCTGTCTGTGTGCCGATGTCTTGACACCTGAGACTTCAGGTTCACCCATCCT
GGTTCTGCCATTCCATTGCAGGGTGGCTTCCCTCCTTTGGGGACTCTTAACGCTTTGGTCTGTTAAAAAA
AAAAAAAAAAAAAAATCCGGGCGTGGTGGCTCACTCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGG
CTGATCATCTGAGGTCAGGGGTTCGAGGCCAGCCCTGACCAACATGGTGAAACCCCGTCTCTACT

Human Arf1 mRNA sequence - var7 (public gi: 14714585) (SEQ ID NO: 331)
CAACGCCTGGCTCGGAGCAGCAGCCTCTGAGGTGTCCCTGGCCAGTGTCCTTCCACCTGTCCACAAGCAT
GGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCATCCTCATGGTGGGC
CTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTGACCACCATTCCCA
CCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGGACGTGGGTGGCCA
GGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTTCGTGGTGGACAGC
AATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAGGACGAGCTCCGGG
ATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGGCCGAGATCACAGA
CAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGCGCCACCAGCGGCGAC
GGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTGAACGCGACCCCCCTCCC
TCTCACTCCTCTTGCCCTCTGCTTTACTCTCATGTGGCAAACGTGCGGCTCGTGGTGTGAGTGCCAGAAG
CTGCCTCCGTGGTTTGGTCACCGTGTGCATCGCACCGTGCTGTAAATGTGGCAGACGCAGCCTGCGGCCA
GGCTTTTTATTTAATGTAAATAGTTTTTGTTTCCAATGAGGCAGTTTCTGGTACTCCTATGCAATATTAC
TCAGCTTTTTTTATTGTAAAAAGAAAAATCAACTCACTGTTCAGTGCTGAGAGGGGATGTAGGCCCATGG
GCACCTGGCCTCCAGGAGTCGCTGTGTTGGGAGAGCCGGCCACGCCCTTGGCTTTAGAGCTGTGTTGAAA
TCCATTTTGGTGGTTGGTTTTTAACCCAAACTCAGTGCATTTTTTAAAATAGTTAAGAATCCAAGTCGAG
AACACTTGAACACACAGAAGGGAGACCCCGCCTAGCATAGATTTGCAGTTACGGCCTGGATGCCAGTCGC
CAGCCCAGCTGTTCCCCTCGGAACATGAGGTGGTGGTGGCGCAGCAGACTGCGATCAATTCTGCATGGT
CACAGTAGAGATCCCCGCAACTCGCTTGTCCTTGGGTCACCCTGCATTCCATAGCCATGTGCTTGTCCCT
GTGCTCCCACGGTTCCCAGGGGCCAGGCTGGGAGCCCACAGCCACCCCACTATGCCGCAGGCCGCCCTAC
CCACCTTCAGGCAGCCTATGGGACGCAGGGCCCCATCTGTCCCTCGGTCGCCGTGTGGCCAGAGTGGGTC
CGTCGTCCCCAACACTCGTGCTCGCTCAGACACTTTGGCAGGATGTCTGGGGCCTCACCAGCAGGAGCGC
GTGCAAGCCGGGCAGGCGGTCCACCTAGACCCACAGCCCCTCGGGAGCACCCCACCTCTGTGTGTGATGT
AGCTTTCTCTCCCTCAGCCTGCAAGGGTCCGATTTGCCATCGAAAAAGACAACCTCTACTTTTTTCTTTT
GTATTTTGATAAACACTGAAGCTGGAGCTGTTAAATTTATCTTGGGGAAACCTCAGAACTGGTCTATTTG
GTGTCGTGGAACCTCTTACTGCTTTCAATACACGATTAGTAATCAACTGTTTTGTATACTTGTTTTCAGT
TTTCATTTCGACAAACAAGCACTGTAATTATAGCTATTAGAATAAAATCTCTTAACTATTAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAA

Human Arf1 mRNA sequence - var8 (public gi: 33872952) (SEQ ID NO: 332)
GTCCAATCAGCTCCGGAACCAGAAGTGAACGCGACCCCCCTCCCTCTCACTCCTCTTGCCCTCTGCTTTA
CTCTCATGTGGCAAACGTGCGGCTCGTGGTGTGAGTGCCAGAAGCTGCCTCCGTGGTTTGGTCACCGTGT
GCATCGCACCGTGCTGTAAATGTGGCAGACGCAGCCTGCGGCCAGGCTTTTTATTTAATGTAAATAGTTT
TTGTTTCCAATGAGGCAGTTTCTGGTACTCCTATGCAATATTACTCAGCTTTTTTTATTGTAAAAAGAAA
AATCAACTCACTGTTCAGTGCTGAGAGGGGATGTAGGCCCATGGGCACCTGGCCTCCAGGAGTCGCTGTG
TTGGGAGAGCCGGCCACGCCCTTGGCTTTAGAGCTGTGTTGAAATCCATTTTGGTGGTTGGTTTTTAACC
CAAACTCAGTGCATTTTTTAAAATAGTTAAGAATCCAAGTCGAGAACACTTGAACACACAGAAGGGAGAC
CCCGCCTAGCATAGATTTGCAGTTACGGCCTGGATGCCAGTCGCCAGCCCAGCTGTTCCCCTCGGGAACA
TGAGGTGGTGGTGGCGCAGCAGACTGCGATCAATTCTGCATGGTCACAGTAGAGATCCCCGCAACTCGCT
TGTCCTTGGGTCACCCTGCATTCCATAGCCATGTGCTTGTCCCTGTGCTCCCACGGTTCCCAGGGGCCAG
GCTGGGAGCCCACAGCCACCCCACTATGCCGCAGGCCGCCCTACCCACCTTCAGGCAGCCTATGGGACGC
AGGGCCCCATCTGTCCCTCGGTCGCCGTGTGGCCAGAGTGGGTCCGTCGTCCCCAACACTCGTGCTCGCT
CAGACACTTTGGCAGGATGTCTGGGGCCTCACCAGCAGGAGCGCGTGCAAGCCGGGCAGGCGGTCCACCT
AGACCCACAGCCCCTCGGGAGCACCCCACCTCTGTGTGTGATGTAGCTTTCTCTCCCTCAGCCTGCAAGG
GTCCGATTTGCCATCGAAAAAGACAACCTCTACTTTTTTCTTTTGTATTTTGATAAACACTGAAGCTGGA
GCTGTTAAATTTATCTTGGGGAAACCTCAGAACTGGTCTATTTGGTGTCGTGGAACCTCTTACTGCTTTC
AATACACGATTAGTAATCAACTGTTTTGTATACTTGTTTTCAGTTTTCATTTCGACAAACAAGCACTGTA
ATTATAGCTATTAGAATAAAATCTCTTAACTATTAAAAAAAAAAAAAAAAAA

Human Arf1 mRNA sequence - var9 (public gi: 15030200) (SEQ ID NO: 333)
GAGCCGCCATCTTGTGGGAGCAAAACCAACGCCTGGCTCGGAGCAGCAGCCTCTGAGGTGTCCCTGGCCA
GTGTCCTTCCACCTGTCCACAAGCATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAA
AAGAAATGCGCATCCTCATGGTGGGCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCT
GGGTGAGATCGTGACCACCATTCCCACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGC
TTCACTGTGTGGGACGTGGGTGGCCAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACAC
AAGGCCTGATCTTCGTGGTGGACAGCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAG
GATGCTGGCCGAGGACGAGCTCCGGGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAAC
GCCATGAATGCGGCCGAGATCACAGACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTC
AGGCCACCTGCGCCACCAGCGGCGACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAA
CCAGAAGTGAACGCGACCCCCCTCCCTCTCACTCCTCTTGCCCTCTGCTTTACTCTCATGTGGCAAACGT
GCGGCTCGTGGTGTGAGTGCCAGAAGCTGCCTCCGTGGTTTGGTCACCGTGTGCATCGCACCGTGCTGTA
AATGTGGCAGACGCAGCCTGCGGCCAGGCTTTTTATTTAATGTAAATAGTTTTTGTTTCCAATGAGGCAG

TTTCTGGTACTCCTATGCAATATTACTCAGCTTTTTTATTGTAAAAAGAAAAATCAACTCACTGTTCAG
TGCTGAGAGGGGATGTAGGCCCATGGGCACCTGGCCTCCAGGAGTCGCTGTGTTGGGAGAGCCGGCCACG
CCCTTGGCTTTAGAGCTGTGTTGAAATCCATTTTGGTGGTTGGTTTTTAACCCAAACTCAGTGCATTTTT
TAAAATAGTTAAGAATCCAAGTCGAGAACACTTGAACACACAGAAGGGAGACCCCGCCTAGCATAGATTT
GCAGTTACGGCCTGGATGCCAGTCGCCAGCCCAGCTGTTCCCCTCGGGAACATGAGGTGGTGGTGGCGCA
GCAGACTGCGATCAATTCTGCATGGTCACAGTAGAGATCCCCGCAACTCGCTTGTCCTTGGGTCACCCTG
CATTCCATAGCCATGTGCTTGTCCCTGTGCTCCCACGGTTCCCAGGGGCCAGGCTGGGAGCCCACAGCCA
CCCCACTATGCCGCAGGCCGCCCTACCCACCTTCAGGCAGCCTATGGGACGCAGGGCCCCATCTGTCCCT
CGGTCGCCGTGTGGCCAGAGTGGGTCCGTCGTCCCCAACACTCGTGCTCGCTCAGACACTTTGGCAGGAT
GTCTGGGGCCTCACCAGCAGGAGCGCGTGCAAGCCGGGCAGGCGGTCCACCTAGACCCACAGCCCCTCGG
GAGCACCCCACCTCTGTGTGTGATGTAGCTTTCTCTCCCTCAGCCTGCAAGGGTCCGATTTGCCATCGAA
AAAGACAACCTCTACTTTTTTCTTTTGTATTTTGATAAACACTGAAGCTGGAGCTGTTAAATTTATCTTG
GGGAAACCTCAGAACTGGTCTATTTGGTGTCGTGGAACCTCTTACTGCTTTCAATACACGATTAGTAATC
AACTGTTTTGTATACTTGTTTTCAGTTTTCATTTCGACAAACAAGCACTGTAATTATAGCTATTAGAATA
AAATCTCTTAACTATTAAAAAAAAAAAAAAAAAA

Human Arf1 mRNA sequence - var10 (public gi: 16553846) (SEQ ID NO: 334)
GTGGGAGCAAAACCAACGCCTGGCTCGGAGCAGCAGCCTCTGAGGTGTCCCTGGCCAGTGTCCTTCCACC
TGTCCACAAGCATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCAT
CCTCATGGTGGGCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTG
ACCACCATTCCCACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGG
ACGTGGGTGGCCAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTT
CGTGGTGGACAGCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAG
GACGAGCTCCGGGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGG
CCGAGATCACAGACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGCGC
CACCAGCGGCGACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTGAACG
CGACCCCCCTCCCTCTCACTCCTCTTGCCCTCTGCTTTACTCTCATGTGGCAAACGTGCGGCTCGTGGTG
TGAGTGCCAGAAGCTGCCTCCGTGGTTTGGTCACCGTGTGCATCGCACCGTGCTGTAAATGTGGCAGACG
CAGCCTGCGGCCAGGCTTTTTATTTAATGTAAATAGTTTTTGTTTCCAATGAGGCAGTTTCTGGTACTCC
TATGCAATATTACTCAGCTTTTTTATTGTAAAAAGAAAAATCAACTCACTGTTCAGTGCTGAGAGGGGA
TGTAGGCCCATGGGCACCTGGCCTCCAGGAGTCGCTGTGTTGGGAGAGCCGGCCACGCCCTTGGCTTTAG
AGCTGTGTTGAAATCCATTTTGGTGGTTGGTTTTTAACCCAAACTCAGTGCATTTTTTAAAATAGTTAAG
AATCCAAGTCGAGAACACTTGAACACACAGAAGGGAGACCCCGCCTAGCATAGATTTGCAGTTACGGCCT
GGATGCCAGTCGCCAGCCCAGCTGTTCCCCTCGGGAACATGAGGTGGTGGTGGCGCAGCAGACTGCGATC
AATTCTGCATGGTCACAGTAGAGATCCCCGCAACTCGCTTGTCCTTGGGTCACCCTGCATTCCATAGCCA
TGTGCTTGTCCCTGTGCTCCCACGGTTCCCAGGGGCCAGGCTGGGAGCCCACAGCCACCCCACTATGCCG
CAGGCCGTCCTACCCACCTTCAGGCAGCCTATGGGACGCAGGGCCCCATCTGTCCCTCGGTCGCCGTGTG
GCCAGAGTGGGTCCGTCGTCCCCAACACTCGTGCTCGCTCAGACACTTTGGCAGGATGTCTGGGGCCTCA
CCAGCAGGAGCGCGTGCAAGCCGGGCAGGCGGTCCACCTAGACCCACAGCCCCTCGGGAGCACCCCACCT
CTGTGTGTGATGTAGCTTTCTCTCCCTCAGCCTGCAAGGGTCCGATTTGCCATCGAAAAAGACAACCTCT
ACTTTTTTCTTTTGTATTTTGATAAACACTGAAGCTGGAGCTGTTAAATTTATCTTGGGGAAACCTCAGA
ACTGGTCTATTTGGTGTCGTGGAACCTCTTACTGCTTTCAATACACGATTAGTAATCAACTGTTTTGTAT
ACTTGTTTTCAGTTTTCATTTCGACAAACAAGCACTGTAATTATAGCTATTAGAATAAAATCTCTTAACT
ATT

Human Arf1 mRNA sequence - var11 (public gi: 16553799) (SEQ ID NO: 335)
AACCAACGCCTGGCTCGGAGCAGCAGCCTCTGAGGTGTCCCTGGCCAGTGTCCTTCCACCTGTCCACAAG
CATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCATCCTCATGGTG
GGCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTGACCACCATTC
CCACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGGACGTGGGTGG
CCAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTTCGTGGTGGAC
AGCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAGGACGAGCTCC
GGGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGGCCGAGATCAC
AGACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGCGCCACCAGCGGC
GACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTGAACGCGACCCCCCT
CCCTCTCACTCCTCTTGCCCTCTGCTTTACTCTCATGTGGCAAACGTGCGGCTCGTGGTGTGAGTGCCAG
AAGCTGCCTCCGTGGTTTGGTCACCGTGTGCATCGCACCGTGCTGTAAATGTGGCAGACGCAGCCTGCGG
CCAGGCTTTTTATTTAATGTAAATAGTTTTTGTTTCCAATGAGGCAGTTTCTGGTACTCCTATGCAATAT
TACTCAGCTTTTTTATTGTAAAAAGAAAAATCAACTCACTGTTCAGTGCTGAGAGGGGATGTAGGCCCA
TGGGCACCTGGCCTCCAGGAGTCGCTGTGTTGGGAGAGCCGGCCACGCCCTTGGCTTTAGAGCTGTGTTG
AAATCCATTTTGGTGGTTGGTTTTTAACCCAAACTCAGTGCATTTTTTAAAATAGTTAAGAATCCAAGTC
GAGAACACTTGAACACACAGAAGGGAGACCCCGCCTAGCATAGATTTGCAGTTACGGCCTGGATGCCAGT

CGCCAGCCCAGCTGTTCCCCTCGGGAACATGAGGTGGTGGTGGCGCAGCAGACTGCGATCAATTCTGCAT
GGTCACAGTAGAGATCCCCGCAACTCGCTTGTCCTTGGGTCACCCTGCATTCCATAGCCATGTGCTTGTC
CCTGTGCTCCCACGGTTCCCAGGGGCCAGGCTGGGAGCCCACAGCCACCCCACTATGCCGCAGGCCGCCC
TACCCACCTTCAGGCAGCCTATGGGACGCAGGGCCCCATCTGTCCCTCGGTCGCCGTGTGGCCAGAGTGG
GTCCGTCGTCCCCAACACTCGTGCTCGCTCAGACACTTTGGCAGGATGTCTGGGGCCTCACCAGCAGGAG
CGCGTGCAAGCCGGGCAGGCGGTCCACCTAGACCCACAGCCCCTCGGGAGCACCCCACCTCTGTGTGTGA
TGTAGCTTTCTCTCCCTCAGCCTGCAAGGGTCCGATTTGCCATCGAAAAAGACAACCTCTACTTTTTCT
TTTGTATTTTGATAAACACTGAAGCTGGAGCTGTTAAATTTATCTTGGGGAAACCTCAGAACTGGTCTAT
TTGGTGTCGTGGAACCTCTTACTGCTTTCAATACACGATTAGTAATC

Human Arf1 mRNA sequence - var12 (public gi: 20147654) (SEQ ID NO: 336)
ATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCATCCTCATGGTGG
GCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTGACCACCATTCC
CACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGGACGTGGGTGGC
CAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTTCGTGGTGGACA
GCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAGGACGAGCTCCG
GGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGGCCGAGATCACA
GACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGCGCCACCAGCGGCG
ACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTGA

Human Arf1 mRNA sequence - var13 (public gi: 178163) (SEQ ID NO: 337)
AAACCAACGCCTGGCTCGGAGCAGCAGCCTCTGAGGTGTCCCTGGCCAGTGTCCTTCCACCTGTCCACAA
GCATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCATCCTCATGGT
GGGCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTGACCACCATT
CCCACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGGACGTGGGTG
GCCAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTTCGTGGTGGA
CAGCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAGGACGAGCTC
CGGGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGGCCGAGATCA
CAGACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGCGCCACCAGCGG
CGACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTGAACGCGACCCCCC
TCCCTCTCACTCCTCTTGCCCTCTGCTTTACTCTCATGTGGCAAACGTGCGGCTCGTGGTGTGAGTGCCA
GAAGCTGCCTCCGTGGTTTGGTCACCGTGTGCATCGCACCGTGCTGTAAATGTGGCAGACGCAGCCTGCG
GCCAGGCTTTTTATTTAATGTAAATAGTTTTTGTTTCCAATGAGGCAGTTTCTGGTACTCCTATGCAATA
TTACTCAGCTTTTTTTATTGTAAAAAGAAAAATCAACTCACTGTTCAGTGCTGAGAGGGGATGTAGGCCC
ATGGGCACCTGGCCTCCAGGAGTCGCTGTGTTGGGAGAGCCGGCCACGCCCTTGGCTTTAGAGCTGTGTT
GAAATCCATTTTGGTGGTTGGTTTTTAACCCAAACTCAGTGCATTTTTTAAAATAGTTAAGAATCCAAGT
CGAGAACACTTGAACACACAGAAGGGAGACCCCGCCTAGCATAGATTTGCAGTTACGGCCTGGATGCCAG
TCGCCAGCCCAGCTGTTCCCCTCGGGAACATGAGGTGGTGGTGGCGCAGCAGACTGCGATCAATTCTGCA
TGGTCACAGTAGAGATCCCCGCAACTCGCTTGTCCTTGGGTCACCCTGCATTCCATAGCCATGTGCTTGT
CCCTGTGCTCCCACGGTTCCCAGGGGCCAGGCTGGGAGCCCACAGCCACCCCACTATGCCGCAGGCCGCC
CTACCCACCTTCAGGCAGCCTATGGGACGCAGGGCCCCATCTGTCCCTCGGTCGCCGTGTGGCCAGAGTG
GGTCCGTCGTCCCCAACACTCGTGCTCGCTCAGACACTTTGGCAGGATGTCTGGGGCCTCACCAGCAGGA
GCGCGTGCAAGCCGGGCAGGCGGTCCACCTAGACCCACAGCCCCTCGGGAGCACCCCACCTCTGTGTGTG
ATGTAGCTTTCTCTCCCTCAGCCTGCAAGGGTCCGATTTGCCATCGAAAAAGACAACCTCTACTTTTTC
TTTTGTATTTTGATAAACACTGAAGCTGGAGCTGTTAAATTTATCTTGGGGAAACCTCAGAACTGGTCTA
TTTGGTGTCGTGGAACCTCTTACTGCTTTCAATACACGATTAGTAATCAACTGTTTTGTATACTTGTTTT
CAGTTTTCATTTCGACAAACAAGCACTGTAATTATAGCTATTAGAATAAAATCTCTTAACTATTT

Human Arf1 mRNA sequence - var14 (public gi: 178982) (SEQ ID NO: 338)
GGGGAAAACCAACGCCTGGCTCGGAGCAGCAGCCTCTGAGGTGTCCCTGGCCAGTGTCCTTCCACCTGTC
CACAAGCATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCATCCTC
ATGGTGGGCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTGACCA
CCATTCCCACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGGACGT
GGGTGGCCAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTTCGTG
GTGGACAGCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAGGACG
AGCTCCGGGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGGCCGA
GATCACAGACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGCGCCACC
AGCGGCGACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTGAACGCGAC
CCCCCTCCCTCTCACTCCTCTTGCCCTCTGCTTTACTCTCATGTGGCAAACGTGCGGCTCGTGGTGTGAG
TGCCAGAAGCTGCCTCCGTGGTTTGGTCACCGTGTGCATCGCACCGTGCTGTAAATGTGGCAGACGCAGC
CTGCGGCCAGGCTTTTTATTTAATGTAAATAGTTTTTGTTTCCAATGAGGCAGTTTCTGGTACTCCTATG
CAATATTACTCAGCTTTTTTTATTGTAAAAAGAAAAATCAACTCACTGTTCAGTGCTGAGAGGGGATGTA

GGCCCATGGGCACCTGGCCTCCAGGAGTCGCTGTGTTGGGAGAGCCGGCCACGCCCTTGGCTTAGAGCTG
TGTTGAAATCCATTTTGGTGGTTGGTTTTAACCCAAACTCAGTGCATTTTTTAAAATAGTTAAGAATCCA
AGTCGAGAACACTTGAACACACAGAAGGGAGACCCCGCCTAGCATAGATTTGCAGTTACGGCCTGGATGC
CAGTCGCCAGCCCAGCTGTTCCCCTCGGGAACATGAGGTGGTGGTGGCGCAGCAGACTGCGATCAATTCT
GCATGGTCACAGTAGAGATCCCCGCAACTCGCTTGTCCTTGGGTCACCCTGCATTCCATAGCCATGTGCT
TGTCCCTGTGCTCCCACGGTTCCCAGGGGCCAGGCTGGGAGCCCACAGCCACCCCACTATGCCGCAGGCC
GCCCTACCCACCTTCAGGCAGCCTATGGGACGCAGGCCCCATCTGTCCCTCGGTCCGCGTGTGGCCAGAG
TGGTCCGTCGTCCCCAACACTCGTGCTCGCTCAGACACTTTGGCAGGATGTCTGGGGCCTCACCAGCAGG
AGCGCGTGCAAGCCGGGCAGGCGGTCCACCTAGACCCACAGCCCCTCGGGAGCACCCCACCTCTGTGTGT
GATGTAGCTTTCTCTCCCTCAGCCTGCAAGGGTCCGATTTGCCATCGAAAAAGACAACCTCTACTTTTTT
CTTTTGTATTTTGATAAACACTGAAGCTGGAGCTGTTAAATTTATCTTGGGGAAACCTCAGAACTGGTCT
ATTTGGTGTCGTAGGAACCTCTTACTGCTTTCAATACACGATTAGTAATCAACTGTTTTGTATACTTGTT
TTCAGTTTTCATTTCGACAAACAAGCACTGTAATTATAGCTATTAGAATAAAATCTCTTAACTATT

Human Arf1 mRNA sequence - var15 (public gi: 3005720) (SEQ ID NO: 339)
AAACCAACGCCTGGCTCGGAGCAGCAGCCTCTGAGGTGTCCCTGGCCAGTGTCCTTCCACCTGTCCACAA
GCATGGGGAACATCTTCGCCAACCTCTTCAAGGGCCTTTTTGGCAAAAAAGAAATGCGCATCCTCATGGT
GGGCCTGGATGCTGCAGGGAAGACCACGATCCTCTACAAGCTTAAGCTGGGTGAGATCGTGACCACCATT
CCCACCATAGGCTTCAACGTGGAAACCGTGGAGTACAAGAACATCAGCTTCACTGTGTGGGACGTGGGTG
GCCAGGACAAGATCCGGCCCCTGTGGCGCCACTACTTCCAGAACACACAAGGCCTGATCTTCGTGGTGGA
CAGCAATGACAGAGAGCGTGTGAACGAGGCCCGTGAGGAGCTCATGAGGATGCTGGCCGAGGACGAGCTC
CGGGATGCTGTCCTCCTGGTGTTCGCCAACAAGCAGGACCTCCCCAACGCCATGAATGCGGCCGAGATCA
CAGACAAGCTGGGGCTGCACTCACTACGCCACAGGAACTGGTACATTCAGGCCACCTGCGCCACCAGCGG
CGACGGGCTCTATGAAGGACTGGACTGGCTGTCCAATCAGCTCCGGAACCAGAAGTGAACGCGACCCCCC
TCCCTCTCACTCCTCTTGCCCTCTGCTTTACTCTCATGTGGCAAACGTGCGGCTCGTGGTGTGAGTGCCA
GAAGCTGCCTCCGTGGTTTGGTCACCGTGTGCATCGCACCGTGCTGTAAATGTGGCAGACGCACCTGCGG
CCAGGCTTTTTATTTAATGTAAATAGTTTTTGTTTCCAATGAGGCAGTTTCTGGTACTCCTATGCAATAT
TACTCAGCTTTTTTTATTGTAAAAAGAAAAATCAACTCACTGTTCAGTGCTGAGAGGGGATGTAGGCCCA
TGGGCACCTGGCCTCCAGGAGTCGCTGTGTTGGGAGAGCCGGCCACGCCCTTGGCTTTAGAGCTGTGTTG
AAATCCATTTTGGTGGTTGGTTTTTAACCCAAACTCAGTGCATTTTTTAAAATAGTTAAGAATCCAAGTC
GAGAACACTTGAACACACAGAAGGGAGACCCCGCCTAGCATAGATTTGCAGTTACGGCCTGGATGCCAGT
CGCCAGCCCAGCTGTTCCCCTCGGGAACATGAGGTGGTGGTGGCGCAGCAGACTGCGATCAATTCTGCAT
GGTCACAGTAGAGATCCCCGCAACTCGCTTGTCCTTGGTCACCCTGCATTCCATAGCCATGTGCTTGTCC
CTGTGCTCCCACGGTTCCCAGGGGCCAGGCTGGGAGCCCACAGCCACCCCACTATGCCGCAGGCCGCCCT
ACCCACCTTCAGGCAGCCTATGGGACGCAGGGCCCCATCTGTCCCTCGGTCGCCGTGTGGCCAGAGTGGG
TCCGTCGTCCCCAACACTCGTGCTCGCTCAGACACTTTGGCAGGATGTCTGGGGCCTCACCAGCAGGAGC
GCGTGCAAGCCGGGCAGGCGGTCCACCTAGACCCACAGCCCCTCGGGAGCACCCCACCTCTGTGTGTGAT
GTAGCTTTCTCTCCCTCAGCCTGCAAGGGTCCGATTTGCCATCGAAAAAGACAACCTCTACTTTTTTCTT
TTGTATTTTGATAAACACTGAAGCTGGAGCTGTTAAATTTATCTTGGGGAAACCTCAGAACTGGTCTATT
TGGTGTCGTGGAACCTCTTACTGCTTTCAATACACGATTAGTAATCAAAAAAAAAAAAAAAAAAAAAAAA
AAA

Human Arf1 protein sequence - var1 (public gi: 3360491) (SEQ ID NO: 223)
MGNIFANLFKGLFGKKEMRILMVGLDAAGKTTILYKLKLGEIVTTIPTIGFNVETVEYKNISFTVWDVGG
QDKIRPLWRHYFQNTQGLIFVVDSNDRERVNEAREELMRMLAEDELRDAVLLVFANKQDLPNAMNAAEIT
DKLGLHSLRHRNWYIQATCATSGDGLYEGLDWLSNQLRNQK

Figure 36 part - 7

PCT/US04/06308

Unigene Name: ARF5 **Unigene ID:** Hs.430657

Human ARF5 mRNA sequence - var1 (public gi: 178986) (SEQ ID NO: 340)
CCAGTTCCAGCCCGCACCCCGCGTCGGTGCCCGCGCCCCTCCCCGGGCCCCGCCATGGCCTCACCGTGT
CCGCGCTCTTTTCGCGGATCTTCGGGAAGAAGCAGATGCGGATTCTCATGGTTGGCTTGGATGCGGCTGG
CAAGACCACAATCCTGTACAAACTGAAGTTGGGGGAGATTGTCACCACCATCCCAACCATAGGCTTCAAT
GTAGAAACAGTGGAATATAAGAACATCTGTTTCACAGTCTGGGACGTGGGAGGCCAGGACAAGATTCGGC
CTCTGTGGCGGCACTACTTCCAGAACACTCAGGGCCTCATCTTTGTGGTGGACAGTAATGACCGGGAGCG
GGTCCAAGAATCTGCTGATGAACTCCAGAAGATGCTGCAGGAGGACGAGCTGCGGGATGCAGTGCTGCTG
GTATTTGCCAACAAGCAGGACATGCCCAACGCCATGCCCGTGAGCGAGCTGACTGACAAGCTGGGGCTAC
AGCACTTACGCAGCCGCACGTGGTATGTCCAGGCCACCTGTGCCACCCAAGGCACAGGTCTGTACGATGG
TCTGGACTGGCTGTCCCACGAGCTGTCAAAGCGCTAACCAGCCAGGGGCAGGCCCCTGATGCCCGGAAGC
TCCTGCGTGCATCCCCGGGATGACCAGACTCCCGGACTCCTCAGGCAGTGCCCTTTCCTCCCACTTTTCC
TCCCCCATAGCCACAGGCCTCTGCTCCTGCTCCTGCCTGCATGTTCTCTCTGTTGTTGGAGCCTGGAGCC
TTGCTCTCTGGGCACAGAGGGGTCCACTCTCCTGCCTGCTGGGACCTATGGAAGGGGCTTCCTGGCCAAG
GCCCCCTCTTCCAGAGGAGGAGCAGGGATCTGGGTTTCCTTTTTTTTTTCTGTTTTGGGTGTACTCTAGG
GGCCAGGTTGGGAGGGGGAAGGTGAGGGCTTCGGGTGGTGCTATAATGTGGCACTGGATCTTGAGTAATA
AATTTGCTGTGGTTTG

Human ARF5 mRNA sequence - var2 (public gi: 21620017) (SEQ ID NO: 341)
CTCCTCCTGCTGCTGCTGCGCCCCATCCCCCCGCGGCCGGCCAGTTCCAGCCCGCACCCCGCGTCGGTGC
CCGCGCCCCTCCCCGGGCTCCGCCATGGGCCTCACCGTGTCCGCGCTCTTTTCGCGGATCTTCGGGAAGA
AGCAGATGCGGATTCTCATGGTTGGCTTGGATGCGGCTGGCAAGACCACAATCCTGTACAAACTGAAGTT
GGGGGAGATTGTCACCACCATCCCAACCATAGGCTTCAATGTAGAAACAGTGGAATATAAGAACATCTGT
TTCACAGTCTGGGACGTGGGAGGCCAGGACAAGATTCGGCCTCTGTGGCGGCACTACTTCCAGAACACTC
AGGGCCTCATCTTTGTGGTGGACAGTAATGACCGGGAGCGGGTCCAAGAATCTGCTGATGAACTCCAGAA
GATGCTGCAGGAGGACGAGCTGCGGGATGCAGTGCTGCTGGTATTTGCCAACAAGCAGGACATGCCCAAC
GCCATGCCCGTGAGCGAGCTGACTGACAAGCTGGGGCTACAGCACTTACGCAGCCGCACGTGGTATGTCC
AGGCCACCTGTGCCACCCAAGGCACAGGTCTGTACGATGGTCTGGACTGGCTGTCCCACGAGCTGTCAAA
GCGCTAACCAGCCAGGGGCAGGCCCCTGATGCCCGGAAGCTCCTGCGTGCATCCCCGGGATGACCAGACT
CCCGGACTCCTCAGGCAGTGCCCTTTCCTCCCACTTTTCCTCCCCCATAGCCACAGGCCTCTGCTCCTGC
TCCTGCCTGCATGTTCTCTCTGTTGTTGGAGCCTGGAGCCTTGCTCTCTGGGCACAGAGGGGTCCACTCT
CCTGCCTGCTGGGACCTATGGAAGGGGCTTCCTGGCCAAGGCCCCCTCTTCCAGAGGAGGAGCAGGGATC
TGGGTTTCCTTTTTTTTTTCTGTTTTGGGTGTACTCTAGGGGCCAGGTTGGGAGGGGGAAGGTGAGGGCT
TCGGGTGGTGCTATAATGTGGCACTGGATCTTGAGTAATAAATTTGCTGTGGTTTGTAAAAAAAAAAAAA
AAAAAAAAAAAAAAAA

Human ARF5 mRNA sequence - var3 (public gi: 12804364) (SEQ ID NO: 342)
CCCGCGTCGGTGCCCGCGCCCCTCCCCGGGCCCCGCCATGGGCCTCACCGTGTCCGCGCTCTTTTCGCGG
ATCTTCGGGAAGAAGCAGATGCGGATTCTCATGGTTGGCTTGGATGCGGCTGGCAAGACCACAATCCTGT
ACAAACTGAAGTTGGGGGAGATTGTCACCACCATCCCAACCATAGGCTTCAATGTAGAAACAGTGGAATA
TAAGAACATCTGTTTCACAGTCTGGGACGTGGGAGGCCAGGACAAGATTCGGCCTCTGTGGCGGCACTAC
TTCCAGAACACTCAGGGCCTCATCTTTGTGGTGGACAGTAATGACCGGGAGCGGGTCCAAGAATCTGCTG
ATGAACTCCAGAAGATGCTGCAGGAGGACGAGCTGCGGGATGCAGTGCTGCTGGTATTTGCCAACAAGCA
GGACATGCCCAACGCCATGCCCGTGAGCGAGCTGACTGACAAGCTGGGGCTACAGCACTTACGCAGCCGC
ACGTGGTATGTCCAGGCCACCTGTGCCACCCAAGGCACAGGTCTGTACGATGGTCTGGACTGGCTGTCCC
ACGAGCTGTCAAAGCGCTAACCAGCCAGGGGCAGGCCCCTGATGCCCGGAAGCTCCTGCGTGCATCCCCG
GATGACCATACTCCCGGACTCCTCAGGCAGTGCCCTTTCCTCCCACTTTTCCTCCCCCATAGCCACAGGC
CTCTGCTCCTGCTCCTGCCTGCATGTTCTCTCTGTTGTTGGAGCCTGGAGCCTTGCTCTCTGGGCACAGA
GGGGTCCACTCTCCTGCCTGCTGGGACCTATGGAAGGGGCTTCCTGGCCAAGGCCCCCTCTTCCAGAGGA
GGAGCAGGGATCTGGGTTTCCTTTTTTTTTTCTGTTTTGGGTGTACTCTAGGGGCCAGGTTGGGAGGGGG
AAGGTGAGGGCTTCGGGTGGTGCTATAATGTGGCACTGGATCTTGAGTAATAAATTTGCTGTGGTTTGAA
AAAAAAAAAAAAAAAAAA

Human ARF5 mRNA sequence - var4 (public gi: 30583012) (SEQ ID NO: 343)
ATGGGCCTCACCGTGTCCGCGCTCTTTTCGCGGATCTTCGGGAAGAAGCAGATGCGGATTCTCATGGTTG
GCTTGGATGCGGCTGGCAAGACCACAATCCTGTACAAACTGAAGTTGGGGGAGATTGTCACCACCATCCC
AACCATAGGCTTCAATGTAGAAACAGTGGAATATAAGAACATCTGTTTCACAGTCTGGGACGTGGGAGGC
CAGGACAAGATTCGGCCTCTGTGGCGGCACTACTTCCAGAACACTCAGGGCCTCATCTTTGTGGTGGACA
GTAATGACCGGGAGCGGGTCCAAGAATCTGCTGATGAACTCCAGAAGATGCTGCAGGAGGACGAGCTGCG
GGATGCAGTGCTGCTGGTATTTGCCAACAAGCAGGACATGCCCAACGCCATGCCCGTGAGCGAGCTGACT

GACAAGCTGGGGCTACAGCACTTACGCAGCCGCACGTGGTATGTCCAGGCCACCTGTGCCACCCAAGGCA
CAGGTCTGTACGATGGTCTGGACTGGCTGTCCCACGAGCTGTCAAAGCGCTAG

Human ARF5 mRNA sequence - var5 (public gi: 6995999) (SEQ ID NO: 344)
CCGCGTCGGTGCCCGCGCCCCTCCCCGGGCCCCGCCATGGGCCTCACCGTGTCCGCGCTCTTTTCGCGGA
TCTTCGGGAAGAAGCAGATGCGGATTCTCATGGTTGGCTTGGATGCGGCTGGCAAGACCACAATCCTGTA
CAAACTGAAGTTGGGGGAGATTGTCACCACCATCCCAACCATAGGCTTCAATGTAGAAACAGTGGAATAT
AAGAACATCTGTTTCACAGTCTGGGACGTGGGAGGCCAGGACAAGATTCGGCCTCTGTGGCGGCACTACT
TCCAGAACACTCAGGGCCTCATCTTTGTGGTGGACAGTAATGACCGGGAGCGGGTCCAAGAATCTGCTGA
TGAACTCCAGAAGATGCTGCAGGAGGACGAGCTGCGGGATGCAGTGCTGCTGGTATTTGCCAACAAGCAG
GACATGCCCAACGCCATGCCCGTGAGCGAGCTGACTGACAAGCTGGGGCTACAGCACTTACGCAGCCGCA
CGTGGTATGTCCAGGCCACCTGTGCCACCCAAGGCACAGGTCTGTACGATGGTCTGGACTGGCTGTCCCA
CGAGCTGTCAAAGCGCTAACCAGCCAGGGGCAGGCCCCTGATGCCCGGAAGCTCCTGCGTGCATCCCCGG
GATGACCAGACTCCCGGACTCCTCAGGCAGTGCCCTTTCCTCCCACTTTCCTCCCCCATAGCCACAGGC
CTCTGCTCCTGCTCCTGCCTGCATGTTCTCTCTGTTGTTGGAGCCTGGAGCCTTGCTCTCTGGGCACAGA
GGGGTCCACTCTCCTGCCTGCTGGGACCTATGGAAGGGGCTTCCTGGCCAAGGCCCCCTCTTCCAGAGGA
GGAGCAGGGATCTGGGTTTCCTTTTTTTTTTCTGTTTTGGGTGTACTCTAGGGGCCAGGTTGGGAGGGGG
AAGGTGAGGGCTTCGGGTGGTGCTATAATGTGGCACTGGATCTTGAGTAATAAATTTGCTGTGGTTTG

Human ARF5 protein sequence - var1 (public gi: 30583013) (SEQ ID NO: 224)
MGLTVSALFSRIFGKKQMRILMVGLDAAGKTTILYKLKLGEIVTTIPTIGFNVETVEYKNICFTVWDVGG
QDKIRPLWRHYFQNTQGLIFVVDSNDRERVQESADELQKMLQEDELRDAVLLVFANKQDMPNAMPVSELT
DKLGLQHLRSRTWYVQATCATQGTGLYDGLDWLSHELSKR

Unigene Name: ATP6V0C **Unigene ID:** Hs.389107

Human ATP6V0C mRNA sequence - var1 (public gi: 33874373) (SEQ ID NO: 345)
GGTATTTAGAGCGCAGCGGCTGACGGGCCGGATCGCCTTCGCCGCCGCCCGCCCGCAAACCTTCGTGCCC
GGCCCGTCCTCGCCCCCGCCTCCGCCACCGCCTCGGCCCGCAGAGCTTGCCCCCTCCCCACCCGCAGACA
TGTCCGAGTCCAAGAGCGGCCCCGAGTATGCTTCGTTTTCGCCGTCATGGGCGCCTCGGCCGCCATGGT
CTTCAGCGCCCTGGGCGCTGCCTATGGCACAGCCAAGAGCGGTACCGGCATTGCGGCCATGTCTGTCATG
CGGCCGGAGCAGATCATGAAGTCCATCATCCCAGTGGTCATGGCTGGCATCATCGCCATCTACGGCCTGG
TGGTGGCAGTCCTCATCGCCAACTCCCTGAATGACGACATCAGCCTCTACAAGAGCTTCCTCCAGCTGGG
CGCCGGCCTGAGCGTGGGCCTGAGCGGCCTGGCAGCCGGCTTTGCCATCGGCATCGTGGGGGACGCTGGC
GTGCGGGGCACCGCCCAGCAGCCCCGACTATTCGTGGGCATGATCCTGATTCTCATCTTCGCCGAGGTGC
TCGGCCTCTACGGTCTCATCGTCGCCCTCATCCTCTCCACAAAGTAGACCCTCTCCGAGCCCACCAGCCA
CAGAATATTATGTAAAGACCACCCCTCCTCATTCCAGAACGAACAGCCTGACACATACGCACGGGGCCGC
CGCCCCCAGTAGTTGGTCTTGTACATGCGCAGTGTCCTAGTGCCCATCGTCTGTTTCCCCGGCCTTGCCC
CCGCCCGCCCCGTGCCGTGGACATCTGGGCCCACTCATCGCCCCTCCAGGCCCCCGGCGCCCCACCCCCT
AGAGTGCTCTGTGTATGCGGATGATTTAGAATTGTCATTTCTCTTTACTGGATGTTTATTTATAAAGATC
TGGCCTGTTCCTGCGTCTGCGGAGCGGCCCTTGTCTCCCAGCTATCTATAACCTTAGCTAGAGTGTCGCC
TTGTGGGTTCCTGTTGCTGAGACTTCCTGGATGGAGCCGCCCTCACCGCCGGGCCCGTGGCCCTGCGCGG
AGCTGTGTCCAATAAAGTTCTTGGATGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human ATP6V0C mRNA sequence - var2 (public gi: 33872390) (SEQ ID NO: 346)
GGCTGACGGGCCGGATCGCCTTCGCCGCCGCCCGCCCGCAAACCTTCGTGCCCGGCCCGTCCTCGCCCCC
GCCTCCGCCACCGCCTCGGCCCGCAGAGCTTGCCCCCTCCCCACCCGCAGACATGTCCGAGTCCAAGAGC
GGCCCCGAGTATGCTTCGTTTTCGCCGTCATGGGCGCCTCGGCCGCCATGGTCTTCAGCGCCCTGGGCG
CTGCCTATGGCACAGCCAAGAGCGGTACCGGCATTGCGGCCATGTCTGTCATGCGGCCGGAGCAGATCAT
GAAGTCCATCATCCCAGTGGTCATGGCTGGCATCATCGCCATCTACGGCCTGGTGGTGGCAGTCCTCATC
GCCAACTCCCTGAATGACGACATCAGCCTCTACAAGAGCTTCCTCCAGCTGGGCGCCGGCCTGAGCGTGG
GCCTGAGCGGCCTGGCAGCCGGCTTTGCCATCGGCATCGTGGGGGACGCTGGCGTGCGGGGCACCGCCCA
GCAGCCCCGACTATTCGTGGGCATGATCCTGATTCTCATCTTCGCCGAGGTGCTCGGCCTCTACGGTCTC
ATCGTCGCCCTCATCCTCTCCACAAAGTAGACCCTCTCCGAGCCCACCAGCCACAGAATATTATGTAAAG
ACCACCCCTCCTCATTCCAGAACGAACAGCCTGACACATACGCACGGGGCCGCCGCCCCCAGTAGTTGGT
CTTGTACATGCGCAGTGTCCTAGTGCCCATCGTCTGTTTCCCCGGCCTTGCCCCCGCCCGCCCCGTGCCG
TGGACATCTGGGCCCACTCATCGCCCCTCCAGGCCCCCGGCGCCCCACCCCCTAGAGTGCTCTGTGTATG
CGGATGATTTAGAATTGTCATTTCTCTTTACTGGATGTTTATTTATAAAGATCTGGCCTGTTCCTGCGTC

TGCGGAGCGGCCCTTGTCTCCCAGCTATCTATAACCTTAGCTAGAGTGTCGCCTTGTGGGTTCCTGTTGC
TGAGACTTCCTGGATGGAGCCGCCCTCACCGCCGGGCCCGTGGCCCTGCGCGGAGCTGTGTCCAATAAAG
TTCTTGGATGTGAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAA

Human ATP6V0C mRNA sequence - var3 (public gi: 33873673) (SEQ ID NO: 347)
CGCCTTCGCCGCCGCCCGCCCGCAAACCTTCGTGCCCGGCCCGTCCTCGCCCCCGCCTCCGCCACCGCCT
CGGCCCGCAGAGCTTGCCCCCTCCCCACCCGCAGACATGTCCGAGTCCAAGAGCGGCCCCGAGTATGCTT
CGTTTTTCGCCGTCATGGGCGCCTCGGCCGCCATGGTCTTCAGCGCCCTGGGCGCTGCCTATGGCACAGC
CAAGAGCGGTACCGGCATTGCGGCCATGTCTGTCATGCGGCCGGAGCAGATCATGAAGTCCATTATCCCA
GTGGTCATGGCTGGCATCATCGCCATCTACGGCCTGGTGGTGGCAGTCCTCATCGCCAACTCCCTGAATG
ACGACATCAGCCTCTACAAGAGCTTCCTCCAGCTGGGCGCCGGCCTGAGCGTGGGCCTGAGCGGCCTGGC
AGCCGGCTTTGCCATCGGCATCGTGGGGGACGCTGGCGTGCGGGGCACCGCCCAGCAGCCCCGACTATTC
GTGGGCATGATCCTGATTCTCATCTTCGCCGAGGTGCTCGGCCTCTACGGTCTCATCGTCGCCCTCATCC
TCTCCACAAAGTAGACCCTCTCCGAGCCCACCAGCCACAGAATATTATGTAAAGACCACCCCTCCTCATT
CCAGAACGAACAGCCTGACACATACGCACGGGGCCGCCGCCCCCAGTAGTTGGTCTTGTACATGCGCAGT
GTCCTAGTGCCCATCGTCTGTTTCCCCGGCCTTGCCCCCGCCCGCCCCGTGCCGTGGACATCTGGGCCCA
CTCATCGCCCCTCCAGGCCCCCGGCGCCCCACCCCCTAGAGTGCTCTGTGTATGCGGATGATTTAGAATT
GTCATTTCTCTTTACTGGATGTTTATTTATAAAGATCTGGCCTGTTCCTGCGTCTGCGGAGCGGCCCTTG
TCTCCCAGCTATCTATAACCTTAGCTAGAGTGTCGCCTTGTGGGTTCCTGTTGCTGAGACTTCCTGGATG
GAGCCGCCCTCACCGCCGGGCCCGTGGCCCTGCGCGGAGCTGTGTCCAATAAAGTTCTTGGATGTGAAAA
AAAAAAAAAAAAAAAAAA

Human ATP6V0C mRNA sequence - var4 (public gi: 33990932) (SEQ ID NO: 348)
GACGGGCCGGATCGCCTTCGCCGCCGCCCGCCCGCAAACCTTCGTGCCCGGCCCGTCCTCGCCCCCGCCT
CCGCCACCGCCTCGGCCCGCAGAGCTTGCCCCCTCCCCACCCGCAGACATGTCCGAGTCCAAGAGCGGCC
CCGAGTATGCTTCGTTTTTCGCCGTCATGGGCGCCTCGGCCGCCATGGTCTTCAGCGCCCTGGGCGCTGC
CTATGGCACAGCCAAGAGCGGTACCGGCATTGCGGCCATGTCTGTCATGCGGCCGGAGCAGATCATGAAG
TCCATCATCCCAGTGGTCATGGCTGGCATCATCGCCATCTACGGCCTGGTGGTGGCAGTCCTCATCGCCA
ACTCCCTGAATGACGACATCAGCCTCTACAAGAGCTTCCTCCAGCTGGGCGCCGGCCTGAGCGTGGGCCT
GAGCGGCCTGGCAGCCGGCTTTGCCATCGGCATCGTGGGGGACGCTGGCGTGCGGGGCACCGCCCAGCAG
CCCCGACTATTCGTGGGCATGATCCTGATTCTCATCTTCGCCGAGGTGCTCGGCCTCTACGGTCTCATCG
TCGCCCTCATCCTCTCCACAAAGTAGACCCTCTCCGAGCCCACCAGCCACAGAATATTATGTAAAGACCA
CCCCTCCTCATTCCAGAACGAACAGCCTGACACATACGCACGGGGCCGCCGCCCCCAGTAGTTGGTCTTG
TACATGCGCAGTGTCCTAGTGCCCATCGTCTGTTTCCCCGGCCTTGCCCCCGCCCGCCCCGTGCCGTGGA
CATCTGGGCCCACTCATCGCCCCTCCAGGCCCCCGGCGCCCCACCCCCTAGAGTGCTCTGTGTATGCGGA
TGATTTAGAATTGTCATTTCTCTTTACTGGATGTTTATTTATAAAGATCTGGCCTGTTCCTGCGTCTGCG
GAGCGGCCCTTGTCTCCCAGCTATCTATAACCTTAGCTAGAGTGTCGCCTTGTGGGTTCCTGTTGCTGAG
ACTTCCTGGATGGAGCCGCCCTCACCGCCGGGCCCGTGGCCCTGCGCGGAGCTGTGTCCAATAAAGTTCT
TGGATGTGAAAAAAAAAAAAAAAAAAA

Human ATP6V0C mRNA sequence - var5 (public gi: 19913436) (SEQ ID NO: 349)
GTTCTGCGGTGCTGGTATTTAGAGCGCAGCGGCTGACGGGCCGGATCGCCTTCGCCGCCGCCCGCCCGCA
AACCTTCGTGCCCGGCCCGTCCTCGCCCCCGCCTCCGCCACCGCCTCGGCCCGCAGAGCTTGCCCCCTCC
CCACCCGCAGACATGTCCGAGTCCAAGAGCGGCCCCGAGTATGCTTCGTTTTTCGCCGTCATGGGCGCCT
CGGCCGCCATGGTCTTCAGCGCCCTGGGCGCTGCCTATGGCACAGCCAAGAGCGGTACCGGCATTGCGGC
CATGTCTGTCATGCGGCCGGAGCAGATCATGAAGTCCATCATCCCAGTGGTCATGGCTGGCATCATCGCC
ATCTACGGCCTGGTGGTGGCAGTCCTCATCGCCAACTCCCTGAATGACGACATCAGCCTCTACAAGAGCT
TCCTCCAGCTGGGCGCCGGCCTGAGCGTGGGCCTGAGCGGCCTGGCAGCCGGCTTTGCCATCGGCATCGT
GGGGGACGCTGGCGTGCGGGGCACCGCCCAGCAGCCCCGACTATTCGTGGGCATGATCCTGATTCTCATC
TTCGCCGAGGTGCTCGGCCTCTACGGTCTCATCGTCGCCCTCATCCTCTCCACAAAGTAGACCCTCTCCG
AGCCCACCAGCCACAGAATATTATGTAAAGACCACCCCTCCTCATTCCAGAACGAACAGCCTGACACATA
CGCACGGGGCCGCCGCCCCCAGTAGTTGGTCTTGTACATGCGCAGTGTCCTAGTGCCCATCGTCTGTTTC
CCCGGCCTTGCCCCCGCCCGCCCCGTGCCGTGGACATCTGGGCCCACTCATCGCCCCTCCAGGCCCCCGG
CGCCCCACCCCCTAGAGTGCTCTGTGTATGCGGATGATTTAGAATTGTCATTTCTCTTTACTGGATGTTT
ATTTATAAAGATCTGGCCTGTTCCTGCGTCTGCGGAGCGGCCCTTGTCTCCCAGCTATCTATAACCTTAG
CTAGAGTGTCGCCTTGTGGGTTCCTGTTGCTGAGACTTCCTGGATGGAGCCGCCCTCACCGCCGGGCCCG
TGGCCCTGCGCGGAGCTGTGTCCAATAAAGTTCTTGGATGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAA

Human ATP6V0C mRNA sequence - var6 (public gi: 34534447) (SEQ ID NO: 350)

TTTATGCTTGTGTTTCTGCAACTGCTTTCTGGCCCCCCACTCTTTCTGTGGCTGCTGAGCCTAGTGCCGC
TCACAGGTCTGCCTTCTGCAGTCTGGTCAGGCTTGGCCTCCGGACTGGAGTCCAGGGTGCTCATGGTATT
CCGCTCCTGGTGGCCATCCCTTTCTTCCCTGTGCTCCTCTTGGTGCCTCCTCCCCCTGCCAGCCACATGA
TTCTTCCTGCTGCCCTCTGTAGAAAAGGGCCTGGCTCACTTCCTGCCTCTGGTGGACTACTGGCCTCACA
GGGTCCACTACTTGGGTTGCTGAGTTCCCTGTATTCAGTCTCCTGCCAACGTGTCTGCCATGCTCTGGTC
TCTTGTGCATACATGATGCAGTTGGATGTGGTCCTGGGCCTGCAGTGGGAGCCCCCTAAAATGCACTGTA
ATTGCTCTATATGCTTGCCAGGGAAAAAATGCACTGTAACCAGGAGTTCAGGACAGGCGCTGGGACAGGC
CCTGGGCCCCAGTCTGCAGGTGCACTGGGTGTTGGCATGGCATGTCTGGGCACCTCCAGGGTGGCGTGGA
GGAGGCCGTGTGGCTCCCTGGCCCAGGTCTCAGCCTCCTTCCTCCCTCTATAGTCACTCCCTGGATACCC
AGCACCGTCGTCTTGGGTGCCTCTGCAGGTGCTATCCAGAGCCCTTGTCTTATTGCCTTGTTTTTCTGTG
ACTCCTCTCTCCCGCCAACTTGGGATACTTGTCTGTGAAGCCCTTCCCCAGCACCCCCTTCTCCGCTCTC
CTGGAGCATGTCTCTGTGCCTGGAGGTCACCGCGCCTGTGTCCTCACCCCTGCTGAGTGCTGGGACACAG
GGTAGGCAAGTTTTGTGGCCCAAATATATCAATAAAATATGAAGAGGAATGGTAGGGGTAGTCCTGGTCC
CTTCCACCTCTGACATATGTAGTCTTCTGCAGGTCAGGCTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGT
GTGTGTGTGTCTGTCAGAGATTCACTCTTGTTGTTTGTTTGAGACGGAGTCTCTCTGTGTCGCCCAGG
CTGGAGTGCAGTGGCGTGATCTTGACTCACTGCAACCTCCACTCCTGGGTTCAGGCGATTCTCCTGCCTC
AGCCTCCCTAGTAACTGGGATGACAGGCATGCGCCACCACTCCTGGCTAATTTTTGTATTTTTAGTAGAG
ACGAGGTTTCACCATGTTACCCAGGCTAATCTCGAACTTCGGATCACCTGAGGTCAGGAGTTGGAGACCA
GCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAAGAAAGTTAGCCAGGTCTGGTGGTGCG
TGCCTGTAATCCCACTTACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGGAGGCAGAGGTTACA
GTGAGCCGAGATCGCGCCACTGCACTCCACCCTGGGCAACAAGAGCGAAAACTGTCTCAAAAAAAAAAAA
AAAAAATTTTTCATTTGAGGTATTCTTCCAGTAGAAGGTTAGTAAGTTTTTAATGAAACCATTAAAAATT
ACACTTCCCAGAAAATAGATGACATCAGTGCCCCTTGCTACTTTCTCAGTCCTCACTATTGCTTTGAGGG
CCCAGGTACTGAAACTGGTTGTCTTGAGTTTTGTGTCAGCTTTTTCTCCAGTCCATTATCCCCCTCCCTT
GCTTCTGAAGCAGTCTAGGTTAAACTAGCCAGGCAGGTAGTTGTGGACTGGTGATTTTCAAAAGCCCCAC
TTTAGAGATCAGGCCACAGCTTTTTATATCGCACAGGACACATCAGCCTGAGCTGCTGCCTCATGCCTGT
TTCCCCAGGAACCTCACTCCTTTGGTAGAACCTTGGGATTTTAGAAATTGTGGCTTTTCCATAACTCATT
TACTCCAACAGTTGAAGTTACACACATTGCTCCCAAATTTGGAAATAGACCACAGTACCTTACCTTTCAT
TCCCCATCTGGCCTTTACCTTCTTTGCTTCAGTGGTTGAAAACAGTTGCCATATTCAAAGTATAGTAGAT
TTCAACCTCACACAAATGACAAGTCCCATTTTACAATCCTAGGAAGGCCCACCAATTTCATTTCACGCGC
CAGGGCGGCTGCAGTTGGAGGCCGAGGGCAGCCCTCTGCTCACTGAATGTCTTGCATGTGCTGACTGCTG
CCCGCAGTGCTGAACATGCCCCACCGCCCAGGCCCAGCACTGCTTGTTGGGTCAG

Human ATP6V0C mRNA sequence - var7 (public gi: 30583148) (SEQ ID NO: 351)
ATGTCCGAGTCCAAGAGCGGCCCCGAGTATGCTTCGTTTTTCGCCGTCATGGGCGCCTCGGCCGCCATGG
TCTTCAGCGCCCTGGGCGCTGCCTATGGCACAGCCAAGAGCGGTACCGGCATTGCGGCCATGTCTGTCAT
GCGGCCGGAGCAGATCATGAAGTCCATCATCCCAGTGGTCATGGCTGGCATCATCGCCATCTACGGCCTG
GTGGTGGCAGTCCTCATCGCCAACTCCCTGAATGACGACATCAGCCTCTACAAGAGCTTCCTCCAGCTGG
GCGCCGGCCTGAGCGTGGGCCTGAGCGGCCTGGCAGCCGGCTTTGCCATCGGCATCGTGGGGGACGCTGG
CGTGCGGGGCACCGCCCAGCAGCCCCGACTATTCGTGGGCATGATCCTGATTCTCATCTTCGCCGAGGTG
CTCGGCCTCTACGGTCTCATCGTCGCCCTCATCCTCTCCACAAAGTAG

Human ATP6V0C protein sequence - var1 (public gi: 30583149) (SEQ ID NO: 225)
MSESKSGPEYASFFAVMGASAAMVFSALGAAYGTAKSGTGIAAMSVMRPEQIMKSIIPVVMAGIIAIYGL
VVAVLIANSLNDDISLYKSFLQLGAGLSVGLSGLAAGFAIGIVGDAGVRGTAQQPRLFVGMILILIFAEV
LGLYGLIVALILSTK

Human ATP6V0C protein sequence - var2 (public gi: 34534448) (SEQ ID NO: 226)
MILPAALCRKGPGSLPASGGLLASQGPLLGLLSSLYSVSCQRVCHALVSCAYMMQLDVVLGLQWEPPKMH
CNCSICLPGKKCTVTRSSGQALGQALGPSLQVHWVLAWHVWAPPGWRGGGRVAPWPRSQPPSSLYSHSLD
TQHRRLGCLCRCYPEPLSYCLVFL

Human ATP6V0C pray sequence - var1 (SEQ ID NO: 352)
CCGCCATGGAGTACCCATACGACGTACCAGATTACGCTCATATGGCCATGGAGGCCAGTGAATTCCACCC
AAGCAGTGGTATCAACGCAGAGTGGCCATTTGGGGGGTCTGCGGTGCTGGTATTTAGAGCGCAGCGGCTG
ACGGGCCGGATCGCCTTCGCCGCCGCCCGCCCGCAAACCTTCGTGCCCGGCCCGTCCTCGCCCCCGCCTC
CGCCACCGCCTCGGCCCGCAGAGCTTGCCCCCTCCCCCCCATGTCGGCCGCCTCGGCCTCTAGAGGGTGG
GCATCGATACGGGATCCATCGAGCTCGAGCTGCAGATGAATCGTAGATACTGAAAAACCCCGCAAGTTCA
CTTCAACTGTGCATTCGTGCA

<u>Unigene Name</u>: CBLB <u>Unigene ID</u>: Hs.3144 <u>Clone ID</u>: 3GD_114

Human CBL-B mRNA sequence - var1 (public gi: 4757919) (SEQ ID NO: 353)
CTGGGTCCTGTGTGTGCCACAGGGGTGGGGTGTCCAGCGAGCGGTCTCCTCCTCCTGCTAGTGCTGCTGC
GGCGTCCCGCGGCCTCCCCGAGTCGGGCGGGAGGGGAGAGCGGGTGTGGATTTGTCTTGACGGTAATTGT
TGCGTTTCCACGTCTCGGAGGCCTGCGCGCTGGGTTGCTCCTTCTTCGGGAGCGAGCTGTTCTCAGCGAT
CCCACTCCCAGCCGGGGCTCCCCACACACACTGGGCTGCGTGCGTGTGGAGTGGGACCCGCGCACACGCG
TGTCTCTGGACAGCTACGGCGCCGAAAGAACTAAAATTCCAGATGGCAAACTCAATGAATGGCAGAAACC
CTGGTGGTCGAGGAGGAAATCCCCGAAAAGGTCGAATTTTGGGTATTATTGATGCTATTCAGGATGCAGT
TGGACCCCCTAAGCAAGCTGCCGCAGATCGCAGGACCGTGGAGAAGACTTGGAAGCTCATGGACAAAGTG
GTAAGACTGTGCCAAAATCCCAAACTTCAGTTGAAAAATAGCCCACCATATATACTTGATATTTTGCCTG
ATACATATCAGCATTTACGACTTATATTGAGTAAATATGATGACAACCAGAAACTTGCCCAACTCAGTGA
GAATGAGTACTTTAAAATCTACATTGATAGCCTTATGAAAAAGTCAAAACGGGCAATAAGACTCTTTAAA
GAAGGCAAGGAGAGAATGTATGAAGAACAGTCACAGGACAGACGAAATCTCACAAAACTGTCCCTTATCT
TCAGTCACATGCTGGCAGAAATCAAAGCAATCTTTCCCAATGGTCAATTCCAGGGAGATAACTTTCGTAT
CACAAAAGCAGATGCTGCTGAATTCTGGAGAAAGTTTTTTGGAGACAAAACTATCGTACCATGGAAAGTA
TTCAGACAGTGCCTTCATGAGGTCCACCAGATTAGCTCTAGCCTGGAAGCAATGGCTCTAAAATCAACAA
TTGATTTAACTTGCAATGATTACATTTCAGTTTTTGAATTTGATATTTTTACCAGGCTGTTTCAGCCTTG
GGGCTCTATTTTGCGGAATTGGAATTTCTTAGCTGTGACACATCCAGGTTACATGGCATTTCTCACATAT
GATGAAGTTAAAGCACGACTACAGAAATATAGCACCAAACCCGGAAGCTATATTTTCCGGTTAAGTTGCA
CTCGATTGGGACAGTGGGCCATTGGCTATGTGACTGGGGATGGGAATATCTTACAGACCATACCTCATAA
CAAGCCCTTATTTCAAGCCCTGATTGATGGCAGCAGGGAAGGATTTTATCTTTATCCTGATGGGAGGAGT
TATAATCCTGATTTAACTGGATTATGTGAACCTACACCTCATGACCATATAAAAGTTACACAGGAACAAT
ATGAATTATATTGTGAAATGGGCTCCACTTTTCAGCTCTGTAAGATTTGTGCAGAGAATGACAAAGATGT
CAAGATTGAGCCTTGTGGGCATTTGATGTGCACCTCTTGCCTTACGGCATGGCAGGAGTCGGATGGTCAG
GGCTGCCCTTTCTGTCGTTGTGAAATAAAAGGAACTGAGCCCATAATCGTGGACCCCTTTGATCCAAGAG
ATGAAGGCTCCAGGTGTTGCAGCATCATTGACCCCTTTGGCATGCCGATGCTAGACTTGGACGACGATGA
TGATCGTGAGGAGTCCTTGATGATGAATCGGTTGGCAAACGTCCGAAAGTGCACTGACAGGCAGAACTCA
CCAGTCACATCACCAGGATCCTCTCCCCTTGCCCAGAGAAGAAAGCCACAGCCTGACCCACTCCAGATCC
CACATCTAAGCCTGCCACCCGTGCCTCCTCGCCTGGATCTAATTCAGAAAGGCATAGTTAGATCTCCCTG
TGGCAGCCCAACAGGTTCACCAAAGTCTTCTCCTTGCATGGTGAGAAAACAAGATAAACCACTCCCAGCA
CCACCTCCTCCCTTAAGAGATCCTCCTCCACCGCCACCTGAAAGACCTCCACCAATCCCACCAGACAATA
GACTGAGTAGACACATCCATCATGTGGAAAGCGTGCCTTCCAGAGACCCGCCAATGCCTCTTGAAGCATG
GTGCCCTCGGGATGTGTTTGGGACTAATCAGCTTGTGGGATGTCGACTCCTAGGGGAGGGCTCTCCAAAA
CCTGGAATCACAGCGAGTTCAAATGTCAATGGAAGGCACAGTAGAGTGGGCTCTGACCCAGTGCTTATGC
GGAAACACAGACGCCATGATTTGCCTTTAGAAGGAGCTAAGGTCTTTTCCAATGGTCACCTTGGAAGTGA
AGAATATGATGTTCCTCCCCGGCTTTCTCCTCCTCCTCCAGTTACCACCCTCCTCCCTAGCATAAAGTGT
ACTGGTCCGTTAGCAAATTCTCTTTCAGAGAAAACAAGAGACCCAGTAGAGGAAGATGATGATGAATACA
AGATTCCTTCATCCCACCCTGTTTCCCTGAATTCACAACCATCTCATTGTCATAATGTAAAACCTCCTGT
TCGGTCCTGTGATAATGGTCACTGTATGCTGAATGGAACACATGGTCCATCTTCAGAGAAGAAATCAAAC
ATCCCTGACTTAAGCATATATTTAAAGGGTACGTATAGAATATAATTTCCTTTGTGATGTACATCTTAAT
GGTCAGAATTTAAAGGCAAAATTTCATGCCATTGTACTGAAAATACATTAAGGTTTTGTGTTATCCTCTA
GGAGATGTTTTTGATTCAGCCTCTGATCCCGTGCCATTACCACCTGCCAGGCCTCCAACTCGGGACAATC
CAAAGCATGGTTCTTCACTCAACAGGACGCCCTCTGATTATGATCTTCTCATCCCTCCATTAGGTTGAAA
CCTTTAAAAAAGTTTTGAACAACCCACCCCTCCTTCTTTTAATTTCAGAATTTTCAGAATTCAGAGTTCA
GTATAACACAGACTCACTGGGTTGTGAATTTGCCTGAAATTTGAATGGGTTCTCCAGGTGCCGGTGACTC
CCAAGTTCACGAGACCATTACTCCATGTAGATGATTAAGGTAGTAGTGTAGTAGTTGGGCATCAGTCAGG
TTTTAAGCAAGTTGTTTTGTCCATACTAAATGTAGTCTAAAAACACATGAGAGCTTTGTGCTCTAGTAGT
TTTGAAGTGATGACTTGAAGTGTTGAGATTTTCTTTAAGTATAATAATTCTTAATAAATATGAACTTGCT
TTTCTTGCAGCATGAGCACCAGTTCCACTTACGCTAATTAAATTATGCAAAATTAAATAGTTGTATGTAG
AGAACTGATAATAAATTCTGTTTTATTCTAATCATTACAACTGTAACACATTCAAAAAAAAAAA

Human CBL-B mRNA sequence - var2 (public gi: 23273908) (SEQ ID NO: 354)
AGCGGAGTGCTGCTGCGGCGTCCCGCGGCCTCCCCGAGTCGGGCGGGAGGGGAGAGCGGGTGTGGATTTG
TCTTGACGGTAATTGTTGCGTTTCCACGTCTCGGAGGCCTGCGCGCTGGGTTGCTCCTTCTTCGGGAGCG
AGCTGTTCTCAGCGATCCCACTCCCAGCCGGGGCTCCCCACACACACTGGGCTGCGTGCGTGTGGAGTGG
GACCCGCGCACACGCGTGTCTCTGGACAGCTACGGCGCCGAAAGAACTAAAATTCCAGATGGCAAACTCA
ATGAATGGCAGAAACCCTGGTGGTCGAGGAGGAAATCCCCGAAAAGGTCGAATTTTGGGTATTATTGATG
CTATTCAGGATGCAGTTGGACCCCCTAAGCAAGCTGCCGCAGATCGCAGGACCGTGGAGAAGACTTGGAA
GCTCATGGACAAAGTGGTAAGACTGTGCCAAAATCCCAAACTTCAGTTGAAAAATAGCCCACCATATATA
CTTGATATTTTGCCTGATACATATCAGCATTTACGACTTATATTGAGTAAATATGATGACAACCAGAAAC
TTGCCCAACTCAGTGAGAATGAGTACTTTAAAATCTACATTGATAGCCTTATGAAAAAGTCAAAACGGGC
AATAAGACTCTTTAAAGAAGGCAAGGAGAGAATGTATGAAGAACAGTCACAGGACAGACGAAATCTCACA

AAACTGTCCCTTATCTTCAGTCACATGCTGGCAGAAATCAAAGCAATCTTTCCCAATGGTCAATTCCAGG
GAGATAACTTTCGTATCACAAAAGCAGATGCTGCTGAATTCTGGAGAAAGTTTTTGGAGACAAAACTAT
CGTACCATGGAAAGTATTCAGACAGTGCCTTCATGAGGTCCACCAGATTAGCTCTGGCCTGGAAGCAATG
GCTCTAAAATCAACAATTGATTTAACTTGCAATGATTACATTTCAGTTTTTGAATTTGATATTTTTACCA
GGCTGTTTCAGCCTTGGGGCTCTATTTTGCGGAATTGGAATTTCTTAGCTGTGACACATCCAGGTTACAT
GGCATTTCTCACATATGATGAAGTTAAAGCACGACTACAGAAATATAGCACCAAACCCGGAAGCTATATT
TTCCGGTTAAGTTGCACTCGATTGGGACAGTGGGCCATTGGCTATGTGACTGGGGATGGGAATATCTTAC
AGACCATACCTCATAACAAGCCCTTATTTCAAGCCCTGATTGATGGCAGCAGGGAAGGATTTTATCTTTA
TCCTGATGGGAGGAGTTATAATCCTGATTTAACTGGATTATGTGAACCTACACCTCATGACCATATAAAA
GTTACACAGGAACAATATGAATTATATTGTGAAATGGGCTCCACTTTTCAGCTCTGTAAGATTTGTGCAG
AGAATGACAAAGATGTCAAGATTGAGCCTTGTGGGCATTTGATGTGCACCTCTTGCCTTACGGCATGGCA
GGAGTCGGATGGTCAGGGCTGCCCTTTCTGTCGTTGTGAAATAAAAGGAACTGAGCCCATAATCGTGGAT
CCCTTTGATCCAAGAGATGAAGGCTCCAGGTGTTGCAGCATCATTGACCCCTTTGGCATGCCGATGCTCG
ACTTGGACGACGATGATGATCGTGAGGAGTCCTTGATGATGAATCGGTTGGCAAACGTCCGAAAGTGCAC
TGACAGGCAGAACTCACCAGTCACATCACCAGGATCCTCTCCCCTTGCCCAGAGAAGAAAGCCACAGCCT
GACCCACTCCAGATCCCACATCTAAGCCTGCCACCCGTGCCTCCTCGCCTGGATCTAATTCAGAAAGGCA
TAGTTAGATCTCCCTGTGGCAGCCCAACGGGTTCACCAAAGTCTTCTCCTTGCATGGTGAGAAAACAAGA
TAAACCACTCCCAGCACCACCTCCTCCCTTAAGAGATCCTCCTCCACCGCCACCTGAAAGACCTCCACCA
ATCCCACCAGACAATAGACTGAGTAGACACATCCATCATGTGGAAAGCGTGCCTTCCAAAGACCCGCCAA
TGCCTCTTGAAGCATGGTGCCCTCGGGATGTGTTTGGGACTAATCAGCTTGTGGGATGTCGACTCCTAGG
GGAGGGCTCTCCAAAACCTGGAATCACAGCGAGTTCAAATGTCAATGGAAGGCACAGTAGAGTGGGCTCT
GACCCAGTGCTTATGCGGAAACACAGACGCCATGATTTGCCTTTAGAAGGAGCTAAGGTCTTTTCCAATG
GTCACCTTGGAAGTGAAGAATATGATGTTCCTCCCCGGCTTTCTCCTCCTCCTCCAGTTACCACCCTCCT
CCCTAGCATAAAGTGTACTGGTCCGTTAGCAAATTCTCTTTCAGAGAAAACAAGAGACCCAGTAGAGGAA
GATGATGATGAATACAAGATTCCTTCATCCCACCCTGTTTCCCTGAATTCACAACCATCTCATTGTCATA
ATGTAAAACCTCCTGTTCGGTCTTGTGATAATGGTCACTGTATGCTGAATGGAACACATGGTCCATCTTC
AGAGAAGAAATCAAACATCCCTGACTTAAGCATATATTTAAAGGGAGATGTTTTTGATTCAGCCTCTGAT
CCCGTGCCATTACCACCTGCCAGGCCTCCAACTCGGGACAATCCAAAGCATGGTTCTTCACTCAACAGGA
CGCCCTCTGATTATGATCTTCTCATCCCTCCATTAGGTGAAGATGCTTTTGATGCCCTCCCTCCATCTCT
CCCACCTCCCCCACCTCCTGCAAGGCATAGTCTCATTGAACATTCAAAACCTCCTGGCTCCAGTAGCCGG
CCATCCTCAGGACAGGATCTTTTTCTTCTTCCTTCAGATCCCTTTGTTGATCTAGCAAGTGGCCAAGTTC
CTTTGCCTCCCGCTAGAAGGTTACCAGGTGAAAATGTCAAAACTAACAGAACATCACAGGACTATGATCA
GCTTCCTTCATGTTCAGATGGTTCACAGGCACCAGCCAGACCCCCTAAACCACGACCGCGCAGGACTGCA
CCAGAAATTCACCACAGAAAACCCCATGGGCCTGAGGCGGCATTGGAAAATGTCGATGCAAAAATTGCAA
AACTCATGGGAGAGGGTTATGCCTTTGAAGAGGTGAAGAGAGCCTTAGAGATAGCCCAGAATAATGTCGA
AGTTGCCCGGAGCATCCTCCGAGAATTTGCCTTCCCTCCTCCAGTATCCCCACGTCTAAATCTATAGCAG
CCAGAACTGTAGACACCAAAATGGAAAGCAATCGATGTATTCCAAGAGTGTGGAAATAAAGAGAACTGAG
ATGGAATTCAAGAGAGAAGTGTCTCCTCCTCGTGTAGCAGCTTGAGAAGAGGCTTGGGAGTGCAGCTTCT
CAAAGGAGACCGATGCTTGCTCAGGATGTCGACAGCTGTGGCTTCCTTGTTTTTGCTAGCCATATTTTTA
AATCAGGGTTGAACTGACAAAAATAATTTAAAGACGTTTACTTCCCTTGAACTTTGAACCTGTGAAATGC
TTTACCTTGTTTACAATTTGGCAAAGTTGCAGTTTGTTCTTGTTTTTAGTTTAGTTTTGTTTTGGTGTTT
TGATACCTGTACTGTGTTCTTCACAGACCCTTTGTAGCGTGGTCAGGTCTGCTGTAACATTTCCCACCAA
CTCTCTTGCTGTCCACATCAACAGCTAAATCATTTATTCATATGGATCTCTACCATCCCCATGCCTTGCC
CAGGTCCAGTTCCATTTCTCTCATTCACAAGATGCTTTGAAGGTTCTGATTTTCAACTGATCAAACTAAT
GCAAAAAAAAAAGTATGTATTCTTCACTACTGAGTTTCTTCTTTGGAAACCATCACTATTGAGAGATGGG
AAAAACCTGAATGTATAAAGCATTTATTTGTCAATAAACTGCCTTTTGTAAGGGGTTTTCACAAAAAAAA
AAAAAAAA

Human CBL-B mRNA sequence - var3 (public gi: 862406) (SEQ ID NO: 355)
CTGGGTCCTGTGTGTGCCACAGGGGTGGGGTGTCCAGCGAGCGGTCTCCTCCTCCTGCTAGTGCTGCTGC
GGCGTCCCGCGGCCTCCCCGAGTCGGGCGGAGGGGAGAGCGGGTGTGGATTTGTCTTGACGGTAATTGT
TGCGTTTCCACGTCTCGGAGGCCTGCGCGCTGGGTTGCTCCTTCTTCGGGAGCGAGCTGTTCTCAGCGAT
CCCACTCCCAGCCGGGGCTCCCCACACACACTGGGCTGCGTGCGTGTGGAGTGGGACCCGCGCACACGCG
TGTCTCTGGACAGCTACGGCGCCGAAAGAACTAAAATTCCAGATGGCAAACTCAATGAATGGCAGAAACC
CTGGTGGTCGAGGAGGAAATCCCCGAAAAGGTCGAATTTTGGGTATTATTGATGCTATTCAGGATGCAGT
TGGACCCCCTAAGCAAGCTGCCGCAGATCGCAGGACCGTGGAGAAGACTTGGAAGCTCATGGACAAAGTG
GTAAGACTGTGCCAAAATCCCAAACTTCAGTTGAAAAATAGCCCACCATATATACTTGATATTTTGCCTG
ATACATATCAGCATTTACGACTTATATTGAGTAAATATGATGACAACCAGAAACTTGCCCAACTCAGTGA
GAATGAGTACTTTAAAATCTACATTGATAGCCTTATGAAAAAGTCAAAACGGGCAATAAGACTCTTTAAA
GAAGGCAAGGAGAGAATGTATGAAGAACAGTCACAGGACAGACGAAATCTCACAAAACTGTCCCTTATCT
TCAGTCACATGCTGGCAGAAATCAAAGCAATCTTTCCCAATGGTCAATTCCAGGGAGATAACTTTCGTAT
CACAAAAGCAGATGCTGCTGAATTCTGGAGAAAGTTTTTGGAGACAAAACTATCGTACCATGGAAAGTA
TTCAGACAGTGCCTTCATGAGGTCCACCAGATTAGCTCTAGCCTGGAAGCAATGGCTCTAAAATCAACAA

TTGATTTAACTTGCAATGATTACATTTCAGTTTTTGAATTTGATATTTTTACCAGGCTGTTTCAGCCTTG
GGGCTCTATTTTGCGGAATTGGAATTTCTTAGCTGTGACACATCCAGGTTACATGGCATTTCTCACATAT
GATGAAGTTAAAGCACGACTACAGAAATATAGCACCAAACCCGGAAGCTATATTTTCCGGTTAAGTTGCA
CTCGATTGGGACAGTGGGCCATTGGCTATGTGACTGGGGATGGGAATATCTTACAGACCATACCTCATAA
CAAGCCCTTATTTCAAGCCCTGATTGATGGCAGCAGGGAAGGATTTTATCTTTATCCTGATGGGAGGAGT
TATAATCCTGATTTAACTGGATTATGTGAACCTACACCTCATGACCATATAAAAGTTACACAGGAACAAT
ATGAATTATATTGTGAAATGGGCTCCACTTTTCAGCTCTGTAAGATTTGTGCAGAGAATGACAAAGATGT
CAAGATTGAGCCTTGTGGGCATTTGATGTGCACCTCTTGCCTTACGGCATGGCAGGAGTCGGATGGTCAG
GGCTGCCCTTTCTGTCGTTGTGAAATAAAAGGAACTGAGCCCATAATCGTGGACCCCTTTGATCCAAGAG
ATGAAGGCTCCAGGTGTTGCAGCATCATTGACCCCTTTGGCATGCCGATGCTAGACTTGGACGACGATGA
TGATCGTGAGGAGTCCTTGATGATGAATCGGTTGGCAAACGTCCGAAAGTGCACTGACAGGCAGAACTCA
CCAGTCACATCACCAGGATCCTCTCCCCTTGCCCAGAGAAGAAAGCCACAGCCTGACCCACTCCAGATCC
CACATCTAAGCCTGCCACCCGTGCCTCCTCGCCTGGATCTAATTCAGAAAGGCATAGTTAGATCTCCCTG
TGGCAGCCCAACAGGTTCACCAAAGTCTTCTCCTTGCATGGTGAGAAAACAAGATAAACCACTCCCAGCA
CCACCTCCTCCCTTAAGAGATCCTCCTCCACCGCCACCTGAAAGACCTCCACCAATCCCACCAGACAATA
GACTGAGTAGACACATCCATCATGTGGAAAGCGTGCCTTCCAGAGACCCGCCAATGCCTCTTGAAGCATG
GTGCCCTCGGGATGTGTTTGGGACTAATCAGCTTGTGGGATGTCGACTCCTAGGGGAGGGCTCTCCAAAA
CCTGGAATCACAGCGAGTTCAAATGTCAATGGAAGGCACAGTAGAGTGGGCTCTGACCCAGTGCTTATGC
GGAAACACAGACGCCATGATTTGCCTTTAGAAGGAGCTAAGGTCTTTTCCAATGGTCACCTTGGAAGTGA
AGAATATGATGTTCCTCCCCGGCTTTCTCCTCCTCCTCCAGTTACCACCCTCCTCCCTAGCATAAAGTGT
ACTGGTCCGTTAGCAAATTCTCTTTCAGAGAAAACAAGAGACCCAGTAGAGGAAGATGATGATGAATACA
AGATTCCTTCATCCCACCCTGTTTCCCTGAATTCACAACCATCTCATTGTCATAATGTAAAACCTCCTGT
TCGGTCCTGTGATAATGGTCACTGTATGCTGAATGGAACACATGGTCCATCTTCAGAGAAGAAATCAAAC
ATCCCTGACTTAAGCATATATTTAAAGGGAGATGTTTTTGATTCAGCCTCTGATCCCGTGCCATTACCAC
CTGCCAGGCCTCCAACTCGGGACAATCCAAAGCATGGTTCTTCACTCAACAGGACGCCCTCTGATTATGA
TCTTCTCATCCCTCCATTAGGTGAAGATGCTTTTGATGCCCTCCCTCCATCTCTCCCACCTCCCCCACCT
CCTGCAAGGCATAGTCTCATTGAACATTCAAAACCTCCTGGCTCCAGTAGCCGGCCATCCTCAGGACAGG
ATCTTTTTCTTCTTCCTTCAGATCCCTTTGTTGATCTAGCAAGTGGCCAAGTTCCTTTGCCTCCTGCTAG
AAGGTTACCAGGTGAAAATGTCAAAACTAACAGAACATCACAGGACTATGATCAGCTTCCTTCATGTTCA
GATGGTTCACAGGCACCAGCCAGACCCCCTAAACCACGACCGCGCAGGACTGCACCAGAAATTCACCACA
GAAAACCCCATGGGCCTGAGGCGGCATTGGAAAATGTCGATGCAAAAATTGCAAAACTCATGGGAGAGGG
TTATGCCTTTGAAGAGGTGAAGAGAGCCTTAGAGATAGCCCAGAATAATGTCGAAGTTGCCCGGAGCATC
CTCCGAGAATTTGCCTTCCCTCCTCCAGTATCCCCACGTCTAAATCTATAGCAGCCAGAACTGTAGACAC
CAAAATGGAAAGCAATCGATGTATTCCAAGAGTGTGGAAATAAAGAGAACTGAGATGGAATTCAAGAGAG
AAGTGTCTCCTCCTCGTGTAGCAGCTTGAGAAGAGGCTTGGGAGTGCAGCTTCTCAAAGGAGACCGATGC
TTGCTCAGGATGTCGACAGCTGTGGCTTCCTTGTTTTTGCTAGCCATATTTTTAAATCAGGGTTGAACTG
ACAAAAATAATTTAAAGACGTTTACTTCCCTTGAACTTTGAACCTGTGAAATGCTTTACCTTGTTTACAA
TTTGGCAAAGTTGCAGTTTGTTCTTGTTTTTAGTTTAGTTTTGTTTTGGTGTTTTGATACCTGTACTGTG
TTCTTCACAGACCCTTTGTAGCGTGGTCAGGTCTGCTGTAACATTTCCCACCAACTCTCTTGCTGTCCAC
ATCAACAGCTAAATCATTTATTCATATGGATCTCTACCATCCCCATGCCTTGCCCAGGTCCAGTTCCATT
TCTCTCATTCACAAGATGCTTTGAAGGTTCTGATTTTCAACTGATCAAACTAATGCAAAAAAAAAAAGTA
TGTATTCTTCACTACTGAGTTTCTTCTTTGGAAACCATCACTATTGAGAGATGGGAAAAACCTGAATGTA
TAAAGCATTTATTTGTCAATAAACTGCCTTTTGTAAGGGGTTTTCACATAAAAAAAAAAAAA

Human CBL-B mRNA sequence - var4 (public gi: 862408) (SEQ ID NO: 356)
CTGGGTCCTGTGTGTGCCACAGGGGTGGGGTGTCCAGCGAGCGGTCTCCTCCTCCTGCTAGTGCTGCTGC
GGCGTCCCGCGGCCTCCCCGAGTCGGGCGGGAGGGGAGAGCGGGTGTGGATTTGTCTTGACGGTAATTGT
TGCGTTTCCACGTCTCGGAGGCCTGCGCGCTGGGTTGCTCCTTCTTCGGGAGCGAGCTGTTCTCAGCGAT
CCCACTCCCAGCCGGGGCTCCCCACACACACTGGGCTGCGTGCGTGTGGAGTGGGACCCGCGCACACGCG
TGTCTCTGGACAGCTACGGCGCCGAAAGAACTAAAATTCCAGATGGCAAACTCAATGAATGGCAGAAACC
CTGGTGGTCGAGGAGGAAATCCCCGAAAAGGTCGAATTTTGGGTATTATTGATGCTATTCAGGATGCAGT
TGGACCCCCTAAGCAAGCTGCCGCAGATCGCAGGACCGTGGAGAAGACTTGGAAGCTCATGGACAAAGTG
GTAAGACTGTGCCAAAATCCCAAACTTCAGTTGAAAAATAGCCCACCATATATACTTGATATTTTGCCTG
ATACATATCAGCATTTACGACTTATATTGAGTAAATATGATGACAACCAGAAACTTGCCCAACTCAGTGA
GAATGAGTACTTTAAAATCTACATTGATAGCCTTATGAAAAAGTCAAAACGGGCAATAAGACTCTTTAAA
GAAGGCAAGGAGAGAATGTATGAAGAACAGTCACAGGACAGACGAAATCTCACAAAACTGTCCCTTATCT
TCAGTCACATGCTGGCAGAAATCAAAGCAATCTTTCCCAATGGTCAATTCCAGGGAGATAACTTTCGTAT
CACAAAAGCAGATGCTGCTGAATTCTGGAGAAAGTTTTTTGGAGACAAAACTATCGTACCATGGAAAGTA
TTCAGACAGTGCCTTCATGAGGTCCACCAGATTAGCTCTAGCCTGGAAGCAATGGCTCTAAAATCAACAA
TTGATTTAACTTGCAATGATTACATTTCAGTTTTTGAATTTGATATTTTTACCAGGCTGTTTCAGCCTTG
GGGCTCTATTTTGCGGAATTGGAATTTCTTAGCTGTGACACATCCAGGTTACATGGCATTTCTCACATAT
GATGAAGTTAAAGCACGACTACAGAAATATAGCACCAAACCCGGAAGCTATATTTTCCGGTTAAGTTGCA
CTCGATTGGGACAGTGGGCCATTGGCTATGTGACTGGGGATGGGAATATCTTACAGACCATACCTCATAA

CAAGCCCTTATTTCAAGCCCTGATTGATGGCAGCAGGGAAGGATTTTATCTTTATCCTGATGGGAGGAGT
TATAATCCTGATTTAACTGGATTATGTGAACCTACACCTCATGACCATATAAAAGTTACACAGGAACAAT
ATGAATTATATTGTGAAATGGGCTCCACTTTTCAGCTCTGTAAGATTTGTGCAGAGAATGACAAAGATGT
CAAGATTGAGCCTTGTGGGCATTTGATGTGCACCTCTTGCCTTACGGCATGGCAGGAGTCGGATGGTCAG
GGCTGCCCTTTCTGTCGTTGTGAAATAAAAGGAACTGAGCCCATAATCGTGGACCCCTTTGATCCAAGAG
ATGAAGGCTCCAGGTGTTGCAGCATCATTGACCCCTTTGGCATGCCGATGCTAGACTTGGACGACGATGA
TGATCGTGAGGAGTCCTTGATGATGAATCGGTTGGCAAACGTCCGAAAGTGCACTGACAGGCAGAACTCA
CCAGTCACATCACCAGGATCCTCTCCCCTTGCCCAGAGAAGAAAGCCACAGCCTGACCCACTCCAGATCC
CACATCTAAGCCTGCCACCCGTGCCTCCTCGCCTGGATCTAATTCAGAAAGGCATAGTTAGATCTCCCTG
TGGCAGCCCAACAGGTTCACCAAAGTCTTCTCCTTGCATGGTGAGAAAACAAGATAAACCACTCCCAGCA
CCACCTCCTCCCTTAAGAGATCCTCCTCCACCGCCACCTGAAAGACCTCCACCAATCCCACCAGACAATA
GACTGAGTAGACACATCCATCATGTGGAAAGCGTGCCTTCCAGAGACCCGCCAATGCCTCTTGAAGCATG
GTGCCCTCGGGATGTGTTTGGGACTAATCAGCTTGTGGGATGTCGACTCCTAGGGGAGGGCTCTCCAAAA
CCTGGAATCACAGCGAGTTCAAATGTCAATGGAAGGCACAGTAGAGTGGGCTCTGACCCAGTGCTTATGC
GGAAACACAGACGCCATGATTTGCCTTTAGAAGGAGCTAAGGTCTTTTCCAATGGTCACCTTGGAAGTGA
AGAATATGATGTTCCTCCCCGGCTTTCTCCTCCTCCTCCAGTTACCACCCTCCTCCCTAGCATAAAGTGT
ACTGGTCCGTTAGCAAATTCTCTTTCAGAGAAAACAAGAGACCCAGTAGAGGAAGATGATGATGAATACA
AGATTCCTTCATCCCACCCTGTTTCCCTGAATTCACAACCATCTCATTGTCATAATGTAAAACCTCCTGT
TCGGTCCTGTGATAATGGTCACTGTATGCTGAATGGAACACATGGTCCATCTTCAGAGAAGAAATCAAAC
ATCCCTGACTTAAGCATATATTTAAAGGGAGATGTTTTTGATTCAGCCTCTGATCCCGTGCCATTACCAC
CTGCCAGGCCTCCAACTCGGGACAATCCAAAGCATGGTTCTTCACTCAACAGGACGCCCTCTGATTATGA
TCTTCTCATCCCTCCATTAGGTTGAAACCTTTAAAAAAGTTTTGAACAACCCACCCCTCCTTCTTTTAAT
TTCAGAATTTTCAGAATTCAGAGTTCAGTATAACACAGACTCACTGGGTTGTGAATTTGCCTGAAATTTG
AATGGGTTCTCCAGGTGCCGGTGACTCCCAAGTTCACGAGACCATTACTCCATGTAGATGATTAAGGTAG
TAGTGTAGTAGTTGGGCATCAGTCAGGTTTTAAGCAAGTTGTTTTGTCCATACTAAATGTAGTCTAAAAA
CACATGAGAGCTTTGTGCTCTAGTAGTTTGAAGTGATGACTTGAAGTGTTGAGATTTTCTTTAAGTATA
ATAATTCTTAATAAATATGAACTTGCTTTTCTTGCAGCATGAGCACCAGTTCCACTTACGCTAATTAAAT
TATGCAAAATTAAATAGTTGTATGTAGAGAACTGATAATAAATTCTGTTTTATTCTAATCATTACAACTG
TAACACATTCAAAAAAAAAAA

Human CBL-B mRNA sequence - var5 (public gi: 862410) (SEQ ID NO: 357)
CTGGGTCCTGTGTGTGCCACAGGGGTGGGGTGTCCAGCGAGCGGTCTCCTCCTCCTGCTAGTGCTGCTGC
GGCGTCCCGCGGCCTCCCCGAGTCGGGCGGAGGGGAGAGCGGGTGTGGATTTGTCTTGACGGTAATTGT
TGCGTTTCCACGTCTCGGAGGCCTGCGCGCTGGGTTGCTCCTTCTTCGGGAGCGAGCTGTTCTCAGCGAT
CCCACTCCCAGCCGGGGCTCCCCACACACACTGGGCTGCGTGCGTGTGGAGTGGACCCGCGCACACGCG
TGTCTCTGGACAGCTACGGCGCCGAAAGAACTAAAATTCCAGATGGCAAACTCAATGAATGGCAGAAACC
CTGGTGGTCGAGGAGGAAATCCCCGAAAAGGTCGAATTTTGGGTATTATTGATGCTATTCAGGATGCAGT
TGGACCCCCTAAGCAAGCTGCCGCAGATCGCAGGACCGTGGAGAAGACTTGGAAGCTCATGGACAAAGTG
GTAAGACTGTGCCAAAATCCCAAACTTCAGTTGAAAAATAGCCCACCATATATACTTGATATTTTGCCTG
ATACATATCAGCATTTACGACTTATATTGAGTAAATATGATGACAACCAGAAACTTGCCCAACTCAGTGA
GAATGAGTACTTTAAAATCTACATTGATAGCCTTATGAAAAAGTCAAAACGGGCAATAAGACTCTTTAAA
GAAGGCAAGGAGAGAATGTATGAAGAACAGTCACAGGACAGACGAAATCTCACAAAACTGTCCCTTATCT
TCAGTCACATGCTGGCAGAAATCAAAGCAATCTTTCCCAATGGTCAATTCCAGGGAGATAACTTTCGTAT
CACAAAAGCAGATGCTGCTGAATTCTGGAGAAAGTTTTTTGGAGACAAAACTATCGTACCATGGAAAGTA
TTCAGACAGTGCCTTCATGAGGTCCACCAGATTAGCTCTAGCCTGGAAGCAATGGCTCTAAAATCAACAA
TTGATTTAACTTGCAATGATTACATTTCAGTTTTTGAATTTGATATTTTTACCAGGCTGTTTCAGCCTTG
GGGCTCTATTTTGCGGAATTGGAATTTCTTAGCTGTGACACATCCAGGTTACATGGCATTTCTCACATAT
GATGAAGTTAAAGCACGACTACAGAAATATAGCACCAAACCCGGAAGCTATATTTTCCGGTTAAGTTGCA
CTCGATTGGGACAGTGGGCCATTGGCTATGTGACTGGGGATGGGAATATCTTACAGACCATACCTCATAA
CAAGCCCTTATTTCAAGCCCTGATTGATGGCAGCAGGGAAGGATTTTATCTTTATCCTGATGGGAGGAGT
TATAATCCTGATTTAACTGGATTATGTGAACCTACACCTCATGACCATATAAAAGTTACACAGGAACAAT
ATGAATTATATTGTGAAATGGGCTCCACTTTTCAGCTCTGTAAGATTTGTGCAGAGAATGACAAAGATGT
CAAGATTGAGCCTTGTGGGCATTTGATGTGCACCTCTTGCCTTACGGCATGGCAGGAGTCGGATGGTCAG
GGCTGCCCTTTCTGTCGTTGTGAAATAAAAGGAACTGAGCCCATAATCGTGGACCCCTTTGATCCAAGAG
ATGAAGGCTCCAGGTGTTGCAGCATCATTGACCCCTTTGGCATGCCGATGCTAGACTTGGACGACGATGA
TGATCGTGAGGAGTCCTTGATGATGAATCGGTTGGCAAACGTCCGAAAGTGCACTGACAGGCAGAACTCA
CCAGTCACATCACCAGGATCCTCTCCCCTTGCCCAGAGAAGAAAGCCACAGCCTGACCCACTCCAGATCC
CACATCTAAGCCTGCCACCCGTGCCTCCTCGCCTGGATCTAATTCAGAAAGGCATAGTTAGATCTCCCTG
TGGCAGCCCAACAGGTTCACCAAAGTCTTCTCCTTGCATGGTGAGAAAACAAGATAAACCACTCCCAGCA
CCACCTCCTCCCTTAAGAGATCCTCCTCCACCGCCACCTGAAAGACCTCCACCAATCCCACCAGACAATA
GACTGAGTAGACACATCCATCATGTGGAAAGCGTGCCTTCCAGAGACCCGCCAATGCCTCTTGAAGCATG
GTGCCCTCGGGATGTGTTTGGGACTAATCAGCTTGTGGGATGTCGACTCCTAGGGGAGGGCTCTCCAAAA
CCTGGAATCACAGCGAGTTCAAATGTCAATGGAAGGCACAGTAGAGTGGGCTCTGACCCAGTGCTTATGC

GGAAACACAGACGCCATGATTTGCCTTTAGAAGGAGCTAAGGTCTTTTCCAATGGTCACCTTGGAAGTGA
AGAATATGATGTTCCTCCCCGGCTTTCTCCTCCTCCTCCAGTTACCACCCTCCTCCCTAGCATAAAGTGT
ACTGGTCCGTTAGCAAATTCTCTTTCAGAGAAAACAAGAGACCCAGTAGAGGAAGATGATGATGAATACA
AGATTCCTTCATCCCACCCTGTTTCCCTGAATTCACAACCATCTCATTGTCATAATGTAAAACCTCCTGT
TCGGTCCTGTGATAATGGTCACTGTATGCTGAATGGAACACATGGTCCATCTTCAGAGAAGAAATCAAAC
ATCCCTGACTTAAGCATATATTTAAAGGGTACGTATAGAATATAATTTCCTTTGTGATGTACATCTTAAT
GGTCAGAATTTAAAGGCAAAATTTCATGCCATTGTACTGAAAATACATTAAGGTTTTGTGTTATCCTCTA
GGAGATGTTTTTGATTCAGCCTCTGATCCCGTGCCATTACCACCTGCCAGGCCTCCAACTCGGGACAATC
CAAAGCATGGTTCTTCACTCAACAGGACGCCCTCTGATTATGATCTTCTCATCCCTCCATTAGGTTGAAA
CCTTTAAAAAAGTTTTGAACAACCCACCCCTCCTTCTTTTAATTTCAGAATTTTCAGAATTCAGAGTTCA
GTATAACACAGACTCACTGGGTTGTGAATTTGCCTGAAATTTGAATGGGTTCTCCAGGTGCCGGTGACTC
CCAAGTTCACGAGACCATTACTCCATGTAGATGATTAAGGTAGTAGTGTAGTAGTTGGGCATCAGTCAGG
TTTTAAGCAAGTTGTTTTGTCCATACTAAATGTAGTCTAAAAACACATGAGAGCTTTGTGCTCTAGTAGT
TTTGAAGTGATGACTTGAAGTGTTGAGATTTTCTTTAAGTATAATAATTCTTAATAAATATGAACTTGCT
TTTCTTGCAGCATGAGCACCAGTTCCACTTACGCTAATTAAATTATGCAAAATTAAATAGTTGTATGTAG
AGAACTGATAATAAATTCTGTTTATTCTAATCATTACAACTGTAACACATTCAAAAAAAAAAA

Human CBL-B mRNA sequence - var6 (public gi: 21753192) (SEQ ID NO: 358)
AGTGCTGCTGCGGCGTCCCGCGGCCTCCCCGAGTCGGGCGGGAGGGGAGAGCGGGTGTGGATTTGTCTTG
ACGGTAATTGTTGCGTTTCCACGTCTCGGAGGCCTGCGCGCTGGGTTGCTCCTTCTTCGGGAGCGAGCTG
TTCTCAGCGATCCCACTCCCAGCCGGGGCTCCCCACACACACACTGGGCTGCGTGCGTGTGGAGTGGGACCC
GCGCACACGCGTGTCTCTGGACAGCTACGGCGCCGAAAGAACTAAAATTCCAGATGGCAAACTCAATGAA
TGGCAGAAACCCTGGTGGTCGAGGAGGAAATCCCCGAAAAGGTCGAATTTTGGGTATTATTGATGCTATT
CAGGATGCAGTTGGACCCCCTAAGCAAGCTGCCGCAGATCGCAAAACCTGGAATCACAGCGAGTTCAAAT
GTCAATGGAAGGCACAGTAGAGTGGGCTCTGACCCAGTGCTTATGCGGAAACACAGACGCCATGATTTGC
CTTTAGAAGGAGCTAAGGTCTTTTCCAATGGTCACCTTGGAAGTGAAGAATATGATGTTCCTCCCCGGCT
TTCTCCTCCTCCTCCAGTTACCACCCTCCTCCCTAGCATAAAGTGTACTGGTCCGTTAGCAAATTCTCTT
TCAGAGAAAACAAGAGACCCAGTAGAGGAAGATGATGATGAATACAAGATTCCTTCATCCCACCCTGTTT
CCCTGAATTCACAACCATCTCATTGTCATAATGTAAAACCTCCTGTTCGGTCTTGTGATAATGGTCACTG
TATGCTGAATGGAACACATGGTCCATCTTCAGAGAAGAAATCAAACATCCCTGACTTAAGCATATATTTA
AAGGGAGATGTTTTTGATTCAGCCTCTGATCCCGTGCCATTACCACCTGCCAGGCCTCCAACTCGGGACA
ATCCAAAGCATGGTTCTTCACTCAACAGGACGCCCTCTGATTATGATCTTCTCATCCCTCCATTAGGTGA
AGATGCTTTTGATGCCCTCCCTCCATCTCTCCCACCTCCCCCACCTCCTGCAAGGCATAGTCTCATTGAA
CATTCAAAACCTCCTGGCTCCAGTAGCCGGCCATCCTCAGGACAGGATCTTTTTCTTCTTCCTTCAGATC
CCTTTGTTGATCTAGCAAGTGGCCAAGTTCCTTTGCCTCCTGCTAGAAGGTTACCAGGTGAAAATGTCAA
AACTAACAGAACATCACAGGACTATGATCAGCTTCCTTCATGTTCAGATGGTTCACAGGCATCAGCCAGA
CCCCCTAAACCACGACCGCGCAGGACTGCACCAGAAATTCACCACAGAAAACCCCATGGGCCTGAGGCGG
CATTGGAAAATGTCGATGCAAAAATTGCAAAACTCATGGGAGAGGGTTATGCCTTTGAAGAGGTGAAGAG
AGCCTTAGAGATAGCCCAGAATAATGTCGAAGTTGCCCGGAGCATCCTCCGAGAATTTGCCTTCCCTCCT
CCAGTATCCCCACGTCTAAATCTATAGCAGCCAGAACTGTAGACACCAAAATGGAAAGCAATCGATGTAT
TCCAAGAGTGTGGAAATAAAGAGAACTGAGATGGAATTCAAGAGAGAAGTGTCTCCTCCTCGTGTAGCAG
CTTGAGAAGAGGCTTGGGAGTGCAGCTTCTCAAAGGAGACCGATGCTTGCTCAGGATGTCGACAGCTGTG
GCTTCCTTGTTTTTGCTAGCCATATTTTTAAATCAGGGTTGAACTGACAAAAATAATTTAAAGACGTTTA
CTTCCCTTGAACTTTGAACCTGTGAAATGCTTTACCTTGTTTACAGTTTGGCAAAGTTGCAGTTTGTTCT
TGTTTTTAGTTTAGTTTTGTTTTGGTGTTTTGTACCTGTACTGTGTTCTTCACAGACCCCTTTGTAGCGTG
GTCAGGTCTGCTGTAACATTTCCCACCAACTCTCTTGCTGTCCACATCAACAGCTAAATCATTTATTCAT
ATGGATCTCTACCATCCCCATGCCTTGCCCAGGTCCAGTTCCATTTCTCTCATTCACAAGATGCTTTGAA
GGTTCTGATTTTCAACTGATCAAACTAATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

Human Cbl-b mRNA sequence - var 7 (SEQ ID NO: 359)
CGTNTTTGGNANNCACTACAGGGGATGTTTAATACACACTCACAATGCGCATGATGTNTATAACTATCTATTCNATGAT
G
TAAGATACCCCACTCAAACCCATAAAAAAGAGCATCTTTAATACGACTCACTATANGGCGAGCGCACGCCATGGCAGGT
A
CCCATACGACGTACCAGATTACGCTCATATGGCCATGGAGGCCAGNGAATTCCACCCAAGCNGTGGTATCAACGCANAG
T
GGACTCTGACCCANTGCTTATGCGGAAACACAGACGCCATGATTTGCCTTTAGAAGGAGCTAAGGTCTCTTCCAATGGT
C
ACCTTGGAAGTGAAGAATATGATGTTCCTCCCCGGCTTTCTCCTCCTCCTCCAGTTACCACCCTNCTCCCTAGCATAAA
G
TGTACTGGTCCGTTAGCAAATTCTCTTTCAGAGAAAACAAGAGACCCAGTAGAGGAAGATGATGATGAATACAAGATTC
C

TTCATCCCACCCTGTTTCCCTGAATTCACAACCATCTCATTGTCATAATGTAAAACCTCCTGTTCGGTCTTGTGATAAT
G
GTCACTGTATGCTGAATGGAACACATGGTCCATCTTCAGAGAAGAAATCAAACATCCCTGACTTAAGCATATATTTAAA
G
GGTGAAGATGCTTTTGATGCCCTCCCTCCATCTCTCCCACCTCCCCCACCTCCTGCAAGGCATAGTCTCATTGAACATT
C
AAAACCTCCTGGCTCCAGTAGCCGGCCATCCTCAGGACAGGATCTTTTTCTTCTTCCTTCAGATCCCTTTGTTGATCTA
G
CAAGTGGCCAAGTTCCTTTGCCTCCCGCTAGAAGGTTACCAGGTGAAAATGTCAAAACTAACAGGACATCACAGGACTA
T
GATCAGCTTCCTTCATGTTCAGATGGTTCACAGGCACCAGCCAGACCCCCTAAACCACGACCGCGCAGGACTGCACCAG
A
AATTCACCACAGAAAACCCCATGGGCCTGAGGCGGCATTGGAAAATGTCGATGCAAAAATTGCAAAACTCATGGGAGAG
G
GTTATGCCTTTGAAGAGGTGAAGAGAGCCTTAGAGATAGCCCAGAATAATGTCGAAGTTGCCCGGAGCATCCTCCGAGA
A
TTTGCCTTCCCTCCTCCAGTATCCCCACGTCTAAATCTATAGCAGCCAGAACTGTAGACACCAAAATGGAAAGCAATCG
A
TGTATTCCAAGAGTGTGGAAATAAAGAGAACTGAGATGGAATTCAAGAGAGAAGTGTCTCCTCCTCGTGTAGCAGCTTG
A
GAAGAGGCTTGGGAGTGCAGCTTCTCAAAGAAAACCGATGCTTGCTCAGGATGTCNACAGCTGNGGNCTNCCTTGTTTT
T
GCTAGCCATTTTTTTAAATNAGGGTTGAACTNGANAAAANTATTTAAAAACGTTTACCTCCCTTGAACTTTGAACCTGG
G
AAAGNC

Human Cbl-b Protein sequence - var 7 (SEQ ID NO: 361)
MRKHRRHDLPLEGAKVSSNGHLGSEEYDVPPRLSPPPPVTTLLPSIKCTGPLANSLSEKTRDPVEEDDDEYKIPSSHPV
S
LNSQPSHCHNVKPPVRSCDNGHCMLNGTHGPSSEKKSNIPDLSIYLKGEDAFDALPPSLPPPPPPARHSLIEHSKPPGS
S
SRPSSGQDLFLLPSDPFVDLASGQVPLPPARRLPGENVKTNRTSQDYDQLPSCSDGSQAPARPPKPRPRRTAPEIHHRK
P
HGPEAALENVDAKIAKLMGEGYAFEEVKRALEIAQNNVEVARSILREFAFPPPVSPRLNL

Human cbl-B clone3Gd114 (partial sequence) (SEQ ID NO: 360)
ACTCTGACCCAGTGCTTATGCGGAAACACAGACGCCATGATTTGCCTTTA
GAAGGAGCTAAGGTCTCTTCCAATGGTCACCTTGGAAGTGAAGAATATGA
TGTTCCTCCCCGGCTTTCTCCTCCTCCTCCAGTTACCACCCTCCTCCCTA
GCATAAAGTGTACTGGTCCGTTAGCAAATTCTCTTTCAGAGAAAACAAGA
GACCCAGTAGAGGAAGATGATGATGAATACAAGATTCCTTCATCCCACCC
TGTTTCCCTGAATTCACAACCATCTCATTGTCATAATGTAAAACCTCCTG
TTCGGTCTTGTGATAATGGTCACTGTATGCTGAATGGAACACATGGTCCA
TCTTCAGAGAAGAAATCAAACATCCCTGACTTAAGCATATATTTAAAGGG
TGAAGATGCTTTTGATGCCCTCCCTCCATCTCTCCCACCTCCCCCACCTC
CTGCAAGGCATAGTCTCATTGAACATTCAAAACCTCCTGGCTCCAGTAGC
CGGCCATCCTCAGGACAGGATCTTTTTCTTCTTCCTTCAGATCCCTTTGT
TGATCTAGCAAGTGGCCAAGTTCCTTTGCCTCCCGCTAGAAGGTTACCAG
GTGAAAATGTCAAAACTAACAGGACATCACAGGACTATGATCAGCTTCCT
TCATGTTCAGATGGTTCACAGGCACCAGCCAGACCCCCTAAACCACGACC
GCGCAGGACTGCACCAGAAATTCACCACAGAAAACCCCATGGGCCTGAGG
CGGCATTGGAAAATGTCGATGCAAAAATTGCAAAACTCATGGGAGAGGGT
TATGCCTTTGAAGAGGTGAAGAGAGCCTTAGAGATAGCCCAGAATAATGT
CGAAGTTGCCCGGAGCATCCTCCGAGAATTTGCCTTCCCTCCTCCAGTAT
CCCCACGTCTAAATCTATAGCAGCCAGAACTGTAGACACCAAAATGGAAA
GCAATCGATGTATTCCAAGAGTGTGGAAATAAAGAGAACTGAGATGGAAT
TCAAGAGAGAAGTGTCTCCTCCTCGTGTAGCAGCTTGAGAAGAGGCTTGG
GAGTGCAGCTTCTCAAAGAAAACCGATGCTTGCTCAGGATGTCGACAGCT
GTGGCTTCCTTGTTTTTGCTAGCCATTTTTTTAAATCAGGGTTGAACTGG
AAAAAATTATTTAAAAACGTTTACCTCCCTTGAACTTTGAACCTGGGAAA

GGC

Human CblB protein in 3Gd114 Translation of cbl-B clone3Gd114 starting at base pair 3 (SEQ ID NO: 398)
SDPVLMRKHRRHDLPLEGAKVSSNGHLGSEEYDVPPRLSPPPPVTTLLPS
IKCTGPLANSLSEKTRDPVEEDDDEYKIPSSHPVSLNSQPSHCHNVKPPV
RSCDNGHCMLNGTHGPSSEKKSNIPDLSIYLKGEDAFDALPPSLPPPPPP
ARHSLIEHSKPPGSSSRPSSGQDLFLLPSDPFVDLASGQVPLPPARRLPG
ENVKTNRTSQDYDQLPSCSDGSQAPARPPKPRPRRTAPEIHHRKPHGPEA

Human CBL-B Protein sequence - var1 (public gi: 4757920) (SEQ ID NO: 227)
MANSMNGRNPGGRGGNPRKGRILGIIDAIQDAVGPPKQAAADRRTVEKTWKLMDKVVRLCQNPKLQLKNS
PPYILDILPDTYQHLRLILSKYDDNQKLAQLSENEYFKIYIDSLMKKSKRAIRLFKEGKERMYEEQSQDR
RNLTKLSLIFSHMLAEIKAIFPNGQFQGDNFRITKADAAEFWRKFFGDKTIVPWKVFRQCLHEVHQISSS
LEAMALKSTIDLTCNDYISVFEFDIFTRLFQPWGSILRNWNFLAVTHPGYMAFLTYDEVKARLQKYSTKP
GSYIFRLSCTRLGQWAIGYVTGDGNILQTIPHNKPLFQALIDGSREGFYLYPDGRSYNPDLTGLCEPTPH
DHIKVTQEQYELYCEMGSTFQLCKICAENDKDVKIEPCGHLMCTSCLTAWQESDGQGCPFCRCEIKGTEP
IIVDPFDPRDEGSRCCSIIDPFGMPMLDLDDDDDREESLMMNRLANVRKCTDRQNSPVTSPGSSPLAQRR
KPQPDPLQIPHLSLPPVPPRLDLIQKGIVRSPCGSPTGSPKSSPCMVRKQDKPLPAPPPPLRDPPPPPPE
RPPPIPPDNRLSRHIHHVESVPSRDPPMPLEAWCPRDVFGTNQLVGCRLLGEGSPKPGITASSNVNGRHS
RVGSDPVLMRKHRRHDLPLEGAKVFSNGHLGSEEYDVPPRLSPPPPVTTLLPSIKCTGPLANSLSEKTRD
PVEEDDDEYKIPSSHPVSLNSQPSHCHNVKPPVRSCDNGHCMLNGTHGPSSEKKSNIPDLSIYLKGTYRI

Human CBL-B Protein sequence - var2 (public gi: 23273909) (SEQ ID NO: 228)
MANSMNGRNPGGRGGNPRKGRILGIIDAIQDAVGPPKQAAADRRTVEKTWKLMDKVVRLCQNPKLQLKNS
PPYILDILPDTYQHLRLILSKYDDNQKLAQLSENEYFKIYIDSLMKKSKRAIRLFKEGKERMYEEQSQDR
RNLTKLSLIFSHMLAEIKAIFPNGQFQGDNFRITKADAAEFWRKFFGDKTIVPWKVFRQCLHEVHQISSG
LEAMALKSTIDLTCNDYISVFEFDIFTRLFQPWGSILRNWNFLAVTHPGYMAFLTYDEVKARLQKYSTKP
GSYIFRLSCTRLGQWAIGYVTGDGNILQTIPHNKPLFQALIDGSREGFYLYPDGRSYNPDLTGLCEPTPH
DHIKVTQEQYELYCEMGSTFQLCKICAENDKDVKIEPCGHLMCTSCLTAWQESDGQGCPFCRCEIKGTEP
IIVDPFDPRDEGSRCCSIIDPFGMPMLDLDDDDDREESLMMNRLANVRKCTDRQNSPVTSPGSSPLAQRR
KPQPDPLQIPHLSLPPVPPRLDLIQKGIVRSPCGSPTGSPKSSPCMVRKQDKPLPAPPPPLRDPPPPPPE
RPPPIPPDNRLSRHIHHVESVPSKDPPMPLEAWCPRDVFGTNQLVGCRLLGEGSPKPGITASSNVNGRHS
RVGSDPVLMRKHRRHDLPLEGAKVFSNGHLGSEEYDVPPRLSPPPPVTTLLPSIKCTGPLANSLSEKTRD
PVEEDDDEYKIPSSHPVSLNSQPSHCHNVKPPVRSCDNGHCMLNGTHGPSSEKKSNIPDLSIYLKGDVFD
SASDPVPLPPARPPTRDNPKHGSSLNRTPSDYDLLIPPLGEDAFDALPPSLPPPPPPARHSLIEHSKPPG
SSSRPSSGQDLFLLPSDPFVDLASGQVPLPPARRLPGENVKTNRTSQDYDQLPSCSDGSQAPARPPKPRP
RRTAPEIHHRKPHGPEAALENVDAKIAKLMGEGYAFEEVKRALEIAQNNVEVARSILREFAFPPPVSPRL
NL

Human CBL-B Protein sequence - var3 (public gi: 862407) (SEQ ID NO: 229)
MANSMNGRNPGGRGGNPRKGRILGIIDAIQDAVGPPKQAAADRRTVEKTWKLMDKVVRLCQNPKLQLKNS
PPYILDILPDTYQHLRLILSKYDDNQKLAQLSENEYFKIYIDSLMKKSKRAIRLFKEGKERMYEEQSQDR
RNLTKLSLIFSHMLAEIKAIFPNGQFQGDNFRITKADAAEFWRKFFGDKTIVPWKVFRQCLHEVHQISSS
LEAMALKSTIDLTCNDYISVFEFDIFTRLFQPWGSILRNWNFLAVTHPGYMAFLTYDEVKARLQKYSTKP
GSYIFRLSCTRLGQWAIGYVTGDGNILQTIPHNKPLFQALIDGSREGFYLYPDGRSYNPDLTGLCEPTPH
DHIKVTQEQYELYCEMGSTFQLCKICAENDKDVKIEPCGHLMCTSCLTAWQESDGQGCPFCRCEIKGTEP
IIVDPFDPRDEGSRCCSIIDPFGMPMLDLDDDDDREESLMMNRLANVRKCTDRQNSPVTSPGSSPLAQRR
KPQPDPLQIPHLSLPPVPPRLDLIQKGIVRSPCGSPTGSPKSSPCMVRKQDKPLPAPPPPLRDPPPPPPE
RPPPIPPDNRLSRHIHHVESVPSRDPPMPLEAWCPRDVFGTNQLVGCRLLGEGSPKPGITASSNVNGRHS
RVGSDPVLMRKHRRHDLPLEGAKVFSNGHLGSEEYDVPPRLSPPPPVTTLLPSIKCTGPLANSLSEKTRD
PVEEDDDEYKIPSSHPVSLNSQPSHCHNVKPPVRSCDNGHCMLNGTHGPSSEKKSNIPDLSIYLKGDVFD
SASDPVPLPPARPPTRDNPKHGSSLNRTPSDYDLLIPPLGEDAFDALPPSLPPPPPPARHSLIEHSKPPG
SSSRPSSGQDLFLLPSDPFVDLASGQVPLPPARRLPGENVKTNRTSQDYDQLPSCSDGSQAPARPPKPRP
RRTAPEIHHRKPHGPEAALENVDAKIAKLMGEGYAFEEVKRALEIAQNNVEVARSILREFAFPPPVSPRL
NL

Human CBL-B Protein sequence - var4 (public gi: 862409) (SEQ ID NO: 230)
MANSMNGRNPGGRGGNPRKGRILGIIDAIQDAVGPPKQAAADRRTVEKTWKLMDKVVRLCQNPKLQLKNS
PPYILDILPDTYQHLRLILSKYDDNQKLAQLSENEYFKIYIDSLMKKSKRAIRLFKEGKERMYEEQSQDR

RNLTKLSLIFSHMLAEIKAIFPNGQFQGDNFRITKADAAEFWRKFFGDKTIVPWKVFRQCLHEVHQISSS
LEAMALKSTIDLTCNDYISVFEFDIFTRLFQPWGSILRNWNFLAVTHPGYMAFLTYDEVKARLQKYSTKP
GSYIFRLSCTRLGQWAIGYVTGDGNILQTIPHNKPLFQALIDGSREGFYLYPDGRSYNPDLTGLCEPTPH
DHIKVTQEQYELYCEMGSTFQLCKICAENDKDVKIEPCGHLMCTSCLTAWQESDGQGCPFCRCEIKGTEP
IIVDPFDPRDEGSRCCSIIDPFGMPMLDLDDDDDREESLMMNRLANVRKCTDRQNSPVTSPGSSPLAQRR
KPQPDPLQIPHLSLPPVPPRLDLIQKGIVRSPCGSPTGSPKSSPCMVRKQDKPLPAPPPPLRDPPPPPPE
RPPPIPPDNRLSRHIHHVESVPSRDPPMPLEAWCPRDVFGTNQLVGCRLLGEGSPKPGITASSNVNGRHS
RVGSDPVLMRKHRRHDLPLEGAKVFSNGHLGSEEYDVPPRLSPPPPVTTLLPSIKCTGPLANSLSEKTRD
PVEEDDDEYKIPSSHPVSLNSQPSHCHNVKPPVRSCDNGHCMLNGTHGPSSEKKSNIPDLSIYLKGDVFD
SASDPVPLPPARPPTRDNPKHGSSLNRTPSDYDLLIPPLG

Unigene Name: CENTB1 **Unigene ID:** Hs.337242

Human CENTB1 mRNA sequence - var1 (public gi: 495679) (SEQ ID NO: 37)
GGGGTGAGAGCTCCTCCTAGGACACCCCTTTCCCCTTGGGGAAAGAATTGTGCCCCCAGGCCCTTCCCCG
CGGAGGTCCCTCTCCTCCTTCCCCCTCATCTCCCCTTCCTGGGACAGAAAGTGCCTCCACCTGCATCCCC
AGGGGCCCGGCCTCCAGGGCCCGCTGGCCCCACAGCAGGCAAGCTGAGATGACGGTCAAGCTGGATTTCG
AGGAGTGTCTCAAGGACTCACCCCGTTTCCGAGCCTCTATTGAGCTGGTGGAAGCCGAAGTGTCAGAATT
GGAGACCCGTCTGGAAAAGCTCCTGAAACTGGGCACTGGTCTCCTGGAAAGTGGGCGCCATTACCTTGCT
GCCAGCCGCGCCTTCGTTGTCGGCATTTGTGACCTGGCCCGCCTGGGTCCACCAGAGCCCATGATGGCGG
AGTGTCTGGAAAAATTCACCGTGAGCCTGAACCACAAGCTGGACAGCCATGCGGAGCTTCTAGATGCCAC
CCAACACACACTGCAGCAGCAGATCCAGACCCTGGTCAAGGAAGGTCTGCGGGGTTTCCGAGAGGCTCGC
CGGGATTTCTGGCGGGGGGCTGAGAGCCTGGAGGCTGCCCTGACCCACAACGCAGAGGTTCCCAGGCGCC
GGGCCCAGGAGGCAGAAGAGGCAGGAGCTGCTTTGAGGACGGCTCGAGCTGGGTACCGGGGACGGGCACT
GGATTATGCCCTGCAGATCAACGTGATTGAGGACAAGAGGAAGTTTGACATCATGGAGTTTGTGCTGCGT
TTGGTGGAGGCCCAGGCTACCCATTTCCAGCAGGGCCATGAGGAGCTGAGCCGGCTGTCCCAGTATCGAA
AGGAGCTGGGCGCCCAGTTGCACCAGCTGGTCTTGAATTCAGCACGAGAGAAGAGGGACATGGAGCAGAG
ACACGTGCTGCTGAAACAGAAGGAGCTGGGTGGGGAGGAGCCAGAACCAAGCTTAAGAGAGGGGCCTGGT
GGCCTGGTGATGGAAGGACATCTCTTCAAACGGGCCAGCAACGCATTTAAGACCTGGAGCAGACGCTGGT
TCACCATTCAGAGCAACCAACTGGTTTACCAGAAGAAGTACAAGGACCCTGTGACTGTGGTGGTGGATGA
CCTTCGTCTCTGCACAGTGAAACTCTGCCCTGACTCAGAAAGGCGGTTCTGCTTTGAGGTGGTGTCCACC
AGCAAGTCCTGCCTCCTCCAGGCTGACTCAGAGCGCCTCCTGCAGCTGTGGGTCAGTGCTGTGCAGAGCA
GCATTGCTTCTGCCTTCAGTCAGGCTCGCCTTGATGACAGCCCCCGGGGTCCAGGCCAGGGCTCAGGACA
CCTGGCCATAGGCTCTGCTGCCACCCTGGGCTCTGGTGGAATGGCCAGGGGAAGGGAGCCTGGGGGAGTC
GGGCACGTGGTGGCCCAGGTCCAGAGTGTGGATGGCAATGCCCAGTGCTGCGACTGCCGGGAGCCAGCCC
CGGAGTGGGCCAGCATCAACCTTGGTGTCACCCTCTGCATTCAGTGTTCCGGCATCCACAGGAGCCTTGG
TGTTCACTTCTCCAAAGTCCGGTCTCTGACCCTTGACTCATGGGAGCCAGAACTAGTGAAGCTCATGTGT
GAGCTGGGAAATGTCATCATCAACCAGATCTATGAGGCCCGCGTGGAGGCCATGGCAGTGAAGAAACCAG
GGCCCAGCTGCTCCCGGCAGGAGAAGGAGGCCTGGATTCACGCTAAATACGTGGAGAAGAAGTTCCTGAC
CAAGCTGCCTGAGATTCGAGGGCGAAGAGGTGGCCGGGGGCGCCCAAGGGGGCAGCCTCCTGTGCCCCCA
AAGCCTTCCATCAGGCCCCGGCCAGGGAGCTTGAGATCCAAGCCAGAGCCCCCCTCTGAGGACCTGGGAA
GCCTGCACCCTGGGGCCCTACTGTTTCGAGCGTCTGGGCATCCTCCATCTCTTCCCACCATGGCTGATGC
CCTTGCCCATGGAGCTGATGTCAACTGGGTCAATGGGGGCCAAGATAATGCCACACCGCTGATCCAGGCC
ACAGCTGCTAATTCTCTTCTGGCCTGTGAGTTTCTCCTCCAGAACGGGGCGAACGTGAACCAAGCGGACA
GTGCGGGCCGGGGCCCGCTGCACCACGCAACCATTCTTGGCCACACGGGGCTCGCCTGCCTGTTCCTGAA
ACGGGGAGCTGATCTGGGGGCTCGAGACTCTGAAGGCAGGGACCCTCTGACCATCGCCATGGAAACAGCC
AACGCTGACATCGTCACCCTGCTACGACTGGCAAAGATGAGGGAGGCTGAAGCGGCCCAGGGGCAGGCAG
GAGATGAGACGTATCTTGACATCTTCCGCGACTTCTCCCTCATGGCGTCAGACGACCCGGAGAAGCTGAG
CCGTCGCAGTCATGACCTCCACACGCTGTGACCCGAGGCCCACGGGGCCCGCGCCTGCCTCCCTTCCCCG
CCACCGGGCCCTCTGCCATTAAAGCCTCCGTGCTTCGCTCTTCC

Human CENTB1 mRNA sequence - var2 (public gi: 17391288) (SEQ ID NO: 38)
GAGCTCCTCCTAGGACACCCCTTTCCCCTTGGGGAAAGGATTGTGCCCCCAGGCCCTTCCCCGCGGAGGT
CCCTCTCCTCCTTCCCCCTCATCTCCCCTTCCTGGGACAGAAAGTGCCTCCACCTGCATCCCCAGGGGCC
CGGCCTCCAGGGCCCGCTGGCCCCACAGCAGGCAAGCTGAGATGACGGTCAAGCTGGATTTCGAGGAGTG
TCTCAAGGACTCACCCCGTTTCCGAGCCTCTATTGAGCTGGTGGAAGCCGAAGTGTCAGAATTGGAGACC
CGTCTGGAAAAGCTCCTGAAACTGGGCACTGGTCTCCTGGAAAGTGGGCGCCATTACCTTGCTGCCAGCC
GCGCCTTCGTTGTCGGCATTTGTGACCTGGCCCGCCTGGGTCCACCAGAGCCCATGATGGCGGAGTGTCT
GGAAAAATTCACCGTGAGCCTGAACCACAAGCTGGACAGCCATGCGGAGCTTCTAGATGCCACCCAACAC
ACACTGCAGCAGCAGATCCAGACCCTGGTCAAGGAAGGTCTGCGGGGTTTCCGAGAGGCTCGCCGGGATT
TCTGGCGGGGGGCTGAGAGCCTGGAGGCTGCCCTGACCCACAACGCAGAGGTTCCCAGGCGCCGGGCCCA
GGAGGCAGAAGAGGCAGGAGCTGCTTTGAGGACGGCTCGAGCTGGGTACCGGGGACGGGCACTGGATTAT

GCCCTGCAGATCAACGTGATTGAGGACAAGAGGAAGTTTGACATCATGGAGTTTGTGCTGCGTTTGGTGG
AGGCCCAGGCTACCCATTTCCAGCAGGGCCATGAGGAGCTGAGCCGGCTGTCCCAGTATCGAAAGGAGCT
GGGCGCCCAGTTGCACCAGCTGGTCTTGAATTCAGCACGAGAGAAGAGGGACATGGAGCAGAGACACGTG
CTGCTGAAACAGAAGGAGCTGGGTGGGGAGGAGCCAGAACCAAGCTTAAGAGAGGGGCCTGGTGGCCTGG
TGATGGAAGGACATCTCTTCAAACGGGCCAGCAACGCATTTAAGACCTGGAGCAGACGCTGGTTCACCAT
TCAGAGCAACCAACTGGTTTACCAGAAGAAGTACAAGGACCCTGTGACTGTGGTGGTGGATGACCTTCGT
CTCTGCACAGTGAAACTCTGCCCTGACTCAGAAAGGCGGTTCTGCTTTGAGGTGGTGTCCACCAGCAAGT
CCTGCCTCCTCCAGGCTGACTCAGAGCGCCTCCTGCAGCTGTGGGTCAGTGCTGTGCAGAGCAGCATTGC
TTCTGCCTTCAGTCAGGCTCGCCTTGATGACAGCCCCCGGGGTCCAGGCCAGGGCTCAGGACACCTGGCC
ATAGGCTCTGCTGCCACCCTGGGCTCTGGTGGAATGGCCAGGGGAAGGGAGCCTGGGGGAGTCGGCACG
TGGTGGCCCAGGTCCAGAGTGTGGATGGCAATGCCCAGTGCTGCGACTGCCGGGAGCCAGCCCCGGAGTG
GGCCAGCATCAACCTTGGTGTCACCCTCTGCATTCAGTGTTCCGGCATCCACAGGAGCCTTGGTGTTCAC
TTCTCCAAAGTCCGGTCTCTGACCCTTGACTCATGGGAGCCAGAACTAGTGAAGCTCATGTGTGAGCTGG
GAAATGTCATCATCAACCAGATCTATGAGGCCCGCGTGGAGGCCATGGCAGTGAAGAAACCAGGGCCCAG
CTGCTCCCGGCAGGAGAAGGAGGCCTGGATTCACGCTAAATACGTGGAGAAGAAGTTCCTGACCAAGCTG
CCTGAGATTCGAGGGCGAAGAGGTGGCCGGGGGCGCCCAAGGGGGCAGCCTCCTGTGCCCCCCAAAGCCTT
CCATCAGGCCCCGGCCAGGGAGCTTGAGATCCAAGCCAGAGCCCCCCTCTGAGGACCTGGGAAGCCTGCA
CCCTGGGGCCCTACTGTTTCGAGCGTCTGGGCATCCTCCATCTCTTCCCACCATGGCTGATGCCCTTGCC
CATGGAGCTGATGTCAACTGGGTCAATGGGGGCCAAGATAATGCCACACCGCTGATCCAGGCCACAGCTG
CTAATTCTCTTCTGGCCTGTGAGTTTCTCCTCCAGAACGGGGCGAACGTGAACCAAGCGGACAGTGCGGG
CCGGGGCCCGCTGCACCACGCAACCATTCTTGGCCACACGGGGCTCGCCTGCCTGTTCCTGAAACGGGGA
GCTGATCTGGGGGCTCGAGACTCTGAAGGCAGGGACCCTCTGACCATCGCCATGGAAACAGCCAACGCTG
ACATCGTCACCCTGCTACGACTGGCAAAGATGAGGGAGGCTGAAGCGGCCCAGGGGCAGGCAGGAGATGA
GACGTATCTTGACATCTTCCGCGACTTCTCCCTCATGGCGTCAGACGACCCGGAGAAGCTGAGCCGTCGC
AGTCATGACCTCCACACGCTGTGACCCGAGGCCCACGGGGCCCGCGCCTGCCTCCCTTCCCCGCCACCGG
GCCCTCTGCCATTAAAGCCTCCGTGCTTCGCTCAAAAAAAAAAAAAAA

Human CENTB1 mRNA sequence - var3 (public gi: 34533014) (SEQ ID NO: 39)
ATGTCAGCGTTGGCTGTTTCCATGGCGATGGTCAGAGGGTCCCTGCCTTCAGAGTCTCGAGCCCCCAGAT
CAGCTCCCCGTTTCAGGAACAGGCAGGCGAGCCTGGAGAGAAGAGCCAGGGTCAGCCGGCCGCCCAACTT
CTCCCAGCCTTCCTCCCCATGCCATCATCCCTACCCCGTGTGGCCAAGAATGGTTGCGTGGTGCAGCGGG
CCCCGGCCCGCACTGTCCGCTTGGTTCACGTTCGCCCCGTTCTGGAGGAGAAACTCACAGGCCAGAAGAG
AATTCTGCATGGAGAAGTCGAGAAGGGGGGTTGAGGGTGGCATCCCTAGTGGTGGATTTCAAGATGTCTT
AGGGTGGCGCCAGTTCAGAGAATGGGAGGGTGGAGTGTGGTAATCAGGAGTGTGGAAGGGGTTACAGCTA
ACTGTAACCAAGCTAGGCTTGGCTCTAGCTCTTTGCATGTATTCATATATAAATCCATAGTACAAGCTTT
TGAGGTATGTTACTATTTTACAGATGAGGCTGAGAGGTTAATAACTTGTTAAAAGTCTCCTGTAGGCCGG
GCACAGTGGCTCACGCCAGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACAGGGTCAGGAGA
TCCAGACCATCCTGGCTAGCACGGTGGAGCCCTATCTCTACTAACAATACAAGAAATTAGCCGGGCATGC
TGGCTGGCGCCTGTGGTCCCAGCTACTCGGGCAGCTGAGGCAGGAGAATGGTGTGAACCCGGGAGGCGGA
GCTTGCAGTGAGCCGAGATCGCACCATTGCACTCCGGCCTGGGGGACGGAGCGAGACTGTCTCAAAAAAA
AAAAAAAAGTCTCCTGTAAGAGGTGAGAGCCTGGGTTCAAACTCAGGTTCTCTGCCTCCAAATCACACAC
TCTTAGCAACCAGTCTCTATTGTTGATCTCTCCCTATGGGTGGAAGCCCTAGGGAACAGGTGGTGGGGAA
AGGAGGTAAGGGCAGGGCCCAGAGTCAGGAGTAGGTGTCAGAGCCCTAGGGTGGGGTGGAGAGGTCAGCA
GGGCTCTTACAGCAGCTGTGGCCTGGATCAGCGGTGTGGCATTATCTTGGCCCCCATTGACCCAGTTGAC
ATCAGCTCCATGGGCAAGGGCATCAGCCATGGTGGGAAGAGATGGAGGATGCCCAGACGCTCGAAACAGT
AGGGCCCCAGGGTGCAGGCTTCCCAGGTCCTCAGAGGGGGGCTCTGTTCGGGGGATTTGGTTCTGTTAGG
GGGAAGCAGCTCCGAGTCTGGGAAGAAAACCCTCAGCAGTGTCCCAATGCTATAATGGACAGGTCTCTT
CTAAATGATGGGGAGCTTGGGACTGTGGAGGGAATAGAGTGATGCAAGTGTGGGTATGTGTAAGTATGCG
TATGCATGTGTACGAGTCCCTAGGGTGTGGGGGAGAGACGGCATCATCACCTCATCTGGTCCAACCACAC
TTGGCCTCAGCTCTCAACCCCTGACGCTCCAGCCAAACCCACCCCCTCTCTCTCTCCTTTTCTTGTGCTG
TTGGCACCCCTTACCCTCCCTGCCCACGCCCAGCCCCACATTCCTTCTCATTCTTAATGTCACACTCCAC
CGTAACCCCTGAAACGGCAGTCCGGTCCCTCCGACATTGTCCAGCGGAAGGCCTGGGCTTCACACTCTGT
GCCTCCCGGCGCTACCTGGCACGATGCCGAGCACACAGCAGATGCTCAATGAATGCCCGACCAACCCTAT
ACCTGGCTTGGATCTCAAGCTCCCTGGCCGGGGCCTGATGGAAGGCTTTGGGGGCACAGGAGGCTGCCCC
CTTGGGCGCCCCCGGCCACCTCTTCGCCCTCGAATCTCAGGCAGCTTGGTCAGGAACTTCTTCTCCACGT
ATTTAGCGTGAATCCAGGCCTCCTTCTCCTGCCTGTGGGAGGGGAGAAGCACGCAGTCTTCCCTCTTCTG
CTCCAGGGGTCCCCCATTCCCCTGGGAGGCTAAACCCCAAGCTCACCGGGAGCAGCTGGGCCCTGGTTTC
TTCACTGCCATGGCCTCCACGCGGGCCTCATAGATCTGGTTGATGATGACATTTCCCAGCTCACACATGA
GCTTCAGGAGGCCCAGGCAGAGGCAGAGACAGGGAAGGTGGGGGTGAGTGACTCCTCAGGGATCACGCCC
CTGCACCGCCATGTCCTTGCCCCACCCCAAGTTCTTGCCCCCAATCTTCACAATACGCTAAGTTACCTTC
ACTAGTTCTGGCTCCCATGAGTCAAGGGTCAGAGACCGGACTTTGGAGAAGTGAACACCAAGGCTCCTGG
AGGGCCAGAGGGGGAGGGTCAGGCCCTGTGCAGGGGGGCAGTGGCCTGGGGAGCTGCTGCTGCTCCTGAA
GACACTGGAGGCAAGGCTGGCATGGGGGCCCGTGCAGAGGTGCTGGCCCAGGAGGCAGGGCAGCTGCGG

CCATGTAACCGCCATGTAGCCTTGACCTGGCCCTGGCAGGACTCTGCCTCGTCACCATTCCTTCTTCCTT
AGGTTTCATTTCAAGGCCCTCATCACTCCAGCCACCTCCCTTCTCTAGTGACACTTGTGACACTTTGGCC
TGGACAACCTCTCCCATGTCACCTCCCTTCCACCACACTGAGGTGGGGGCGAGGGCCTTAGATACTTGC
TAAGGCCTCATGACCGTTTCTCTGCCTAGTCTTCACTGGCTCCCCCACCCTCAGCAGCCTTGACCCCACA
CTTCTTCCAACCAAGCCAACAAATTCTGGGTATCCCCCAATTCTGGCCAGACTAGGACACAGAGGGGCTA
GGCCCGCCTGGGTCCAACTGGCACCCCAGAGGCTTGGGCCCAGGCCTGGTACCCAGTGACAAAGCCAGAA
GCTAAGAGAGGAAGCCAGGACAGGGAAGGAAGAGGGGCCGGTGTGATGCGCTCTGTATTGGAGCCGCACT
GTGGCCCGAAGGAGTGGGGCTCCCGCATGGGCCTTGTGGAGTAACCTGTGGATGCCGGAACACTGAATGC
AGAGGGTGACACCAAGGTTGATGCTGGCCCACTCCGGGGCTGGCTCCCGGCAGTCGCAGCACTGGGCATT
GCCATCCACACTCTGGACCTGGGCCACCACGTGCCCGACTCCCCCAGGCTCCCTTCCCCTGGCCATTCCA
CCAGAGCCCAGGGTGGCAGCAGAGCCTATGGCCAGGTGTCCTGAGCCCTGGGGAGAGAGGGGAAGAAAG
GGTGGCCAAGGGGCCTAGGGTAAAGGGTGCCCCATCTCCACAGGCAGCCTGGCTCCGCACCCCCAGGTTA
AGGTACCTGGCCTGGACCCCGGGGGCTGTCATCAAGGCGAGCCTGACTGAAGGCAGAAGCAATGCTGCTC
TGCACAGCACTGACCCACAGCTGCAGGAGGCGCTCTGAGTCAGCCTGGAGGAGGCAGGACCTAGGTAGGA
GGGTGAGGGAGATGGCAGAGGGGTCTGAGGCCTGGGAAGCAAAGTGGCAGCATGGGCAGACTGACATTCA
GCCAGTATTCAACCAGTTCCAGTTGCATTGAAAGACTTCTGTACCAGTTGGTAATATTCTCCTAAATATC
CCCCATCACCCTGTACCCTCTTCCACAATGGCCCCCCAGTCCAGCCGCCAAAGAATTAAATTAAAGTCTG
GAGCTGCATGGGGGGCTTCCATTGTGGTGGGCCCTGCCTTTCAGATTGGCAGTTGTTTAGATATATTAGA
GTATCACCCCTGGGGATTGCACTCACTTGCTGGTGGACACCACCTCAAAGCAGAACCGCCTTTCTGAGTC
AGGGCAGAGTTTCACTGTGCAGAGACGAAGGTCATCCACCACCACAGTCACAGGGTCCTGGCAGGATAAG
GTGATAAGGGGCCAGATGTCCAGCTGCAGGCAAGAGCTGAGTCTCCCTGGGGCCCAGGCATCCAGGACCC
AGGTCCACTCACCTTGTACTTCTTCTGGTAAACCAGTTGGTTGCTCTGAATGGTGAACCAGCGTCTGTAA
GAGAAGGAAATCATTACAGACATAGGCAGCTTTAGGATGAGGGACGGAAGAGAGGCTGTGCTTTTTGCCC
ATGAGGATCTTACTGAGAGGACAGACACCTGGGCTGACTGTTCCACGAGACATTCCAGAGAAGGGTGGAC
AATTGTGCAGATTGGAACATCTAAAGGATGCTATTCCTATCTTGGACAACCCAGATTTCATATAGTTATG
AAGACAACTTTCCAGCAGATGGCAGTAAAATTCTTTTTCTAATAAAATGTCTATTGCTACAATTTAAAAA
ATACTATTTAGGCTGGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCTGATGGGGGTGGTGGATCGCC
CGAGGTCAGGAGTTTGAGACCACCCTGACCAATATGGTGAAACTCCGTCTCTACTAAAAATACAAAAATT
AGCCAGGCGTGGTGGCAGGCGGCTATAATCCCACCTACTTGGGAGGCTGAGGCGGGAGAATCGCTTGAAC
CCAGGAAGCTGAGGTTGCAGTGAGCTGGGATCGCACCACTGTGCTGCAGCCTGCGCAACATAGCGAGGCT
CCATCAAAAAAGAAAAAAAAAAGAAAAAGAAAAAAAGAAAAGAAAGAATCTTGGGGGCCAGGTACAGTGG
CTCACGCCTGTAGTCCCAGCAAGTTGGGAGGCCGAGGCGGGTGGATTGCTTGATGTCAGGAGTTTGCAAC
CAGCCTGGGCAACATGGTGAAACCCTGTTTCTACCAAAAATACAAAAATTAGCCGAGCGTGATGGCACGC
GCCTGTGGTCCCAGCTGTTTAGGATGCTGAGGAGGGAGGATCACTTGAACTCAGGGGATAGAGGTTGCAG
TGAGCCGAGACTGCGCCACTGCACTGCAGGCTGGGCAACAGAGTGACACCCCATCTCAAAAAAAAAACAG

Human CENTB1 mRNA sequence - var4 (public gi: 32879918) (SEQ ID NO: 40)
ATGACGGTCAAGCTGGATTTCGAGGAGTGTCTCAAGGACTCACCCCGTTTCCGAGCCTCTATTGAGCTGG
TGGAAGCCGAAGTGTCAGAATTGGAGACCCGTCTGGAAAAGCTCCTGAAACTGGGCACTGGTCTCCTGGA
AAGTGGGCGCCATTACCTTGCTGCCAGCCGCGCCTTCGTTGTCGGCATTTGTGACCTGGCCCGCCTGGGT
CCACCAGAGCCCATGATGGCGGAGTGTCTGGAAAAATTCACCGTGAGCCTGAACCACAAGCTGGACAGCC
ATGCGGAGCTTCTAGATGCCACCCAACACACACTGCAGCAGCAGATCCAGACCCTGGTCAAGGAAGGTCT
GCGGGGTTTCCGAGAGGCTCGCCGGGATTTCTGGCGGGGGGCTGAGAGCCTGGAGGCTGCCCTGACCCAC
AACGCAGAGGTTCCCAGGCGCCGGGCCCAGGAGGCAGAAGAGGCAGGAGCTGCTTTGAGGACGGCTCGAG
CTGGGTACCGGGGACGGGCACTGGATTATGCCCTGCAGATCAACGTGATTGAGGACAAGAGGAAGTTTGA
CATCATGGAGTTTGTGCTGCGTTTGGTGGAGGCCCAGGCTACCCATTTCCAGCAGGGCCATGAGGAGCTG
AGCCGGCTGTCCCAGTATCGAAAGGAGCTGGGCGCCCAGTTGCACCAGCTGGTCTTGAATTCAGCACGAG
AGAAGAGGGACATGGAGCAGAGACACGTGCTGCTGAAACAGAAGGAGCTGGGTGGGGAGGAGCCAGAACC
AAGCTTAAGAGAGGGGCCTGGTGGCCTGGTGATGGAAGGACATCTCTTCAAACGGGCCAGCAACGCATTT
AAGACCTGGAGCAGACGCTGGTTCACCATTCAGAGCAACCAACTGGTTTACCAGAAGAAGTACAAGGACC
CTGTGACTGTGGTGGTGGATGACCTTCGTCTCTGCACAGTGAAACTCTGCCCTGACTCAGAAAGGCGGTT
CTGCTTTGAGGTGGTGTCCACCAGCAAGTCCTGCCTCCTCCAGGCTGACTCAGAGCGCCTCCTGCAGCTG
TGGGTCAGTGCTGTGCAGAGCAGCATTGCTTCTGCCTTCAGTCAGGCTCGCCTTGATGACAGCCCCCGGG
GTCCAGGCCAGGGCTCAGGACACCTGGCCATAGGCTCTGCTGCCACCCTGGGCTCTGGTGGAATGGCCAG
GGGAAGGGAGCCTGGGGGAGTCGGGCACGTGGTGGCCCAGGTCCAGAGTGTGGATGGCAATGCCCAGTGC
TGCGACTGCCGGGAGCCAGCCCCGGAGTGGGCCAGCATCAACCTTGGTGTCACCCTCTGCATTCAGTGTT
CCGGCATCCACAGGAGCCTTGGTGTTCACTTCTCCAAAGTCCGGTCTCTGACCCTTGACTCATGGGAGCC
AGAACTAGTGAAGCTCATGTGTGAGCTGGGAAATGTCATCATCAACCAGATCTATGAGGCCCGCGTGGAG
GCCATGGCAGTGAAGAAACCAGGGCCCAGCTGCTCCCGGCAGGAGAAGGAGGCCTGGATTCACGCTAAAT
ACGTGGAGAAGAAGTTCCTGACCAAGCTGCCTGAGATTCGAGGGCGAAGAGGTGGCCGGGGGCGCCCAAG
GGGGCAGCCTCCTGTGCCCCCAAAGCCTTCCATCAGGCCCCGGCCAGGGAGCTTGAGATCCAAGCCAGAG
CCCCCCTCTGAGGACCTGGGAAGCCTGCACCCTGGGGCCCTACTGTTTCGAGCGTCTGGGCATCCTCCAT
CTCTTCCCACCATGGCTGATGCCCTTGCCCATGGAGCTGATGTCAACTGGGTCAATGGGGGCCAAGATAA

TGCCACACCGCTGATCCAGGCCACAGCTGCTAATTCTCTTCTGGCCTGTGAGTTTCTCCTCCAGAACGGG
GCGAACGTGAACCAAGCGGACAGTGCGGGCCGGGGCCCGCTGCACCACGCAACCATTCTTGGCCACACGG
GGCTCGCCTGCCTGTTCCTGAAACGGGGAGCTGATCTGGGGGCTCGAGACTCTGAAGGCAGGGACCCTCT
GACCATCGCCATGGAAACAGCCAACGCTGACATCGTCACCCTGCTACGACTGGCAAAGATGAGGGAGGCT
GAAGCGGCCCAGGGGCAGGCAGGAGATGAGACGTATCTTGACATCTTCCGCGACTTCTCCCTCATGGCGT
CAGACGACCCGGAGAAGCTGAGCCGTCGCAGTCATGACCTCCACACGCTGTAG

Human CENTB1 protein sequence - var1 (public gi: 32879919) (SEQ ID NO: 231)
MTVKLDFEECLKDSPRFRASIELVEAEVSELETRLEKLLKLGTGLLESGRHYLAASRAFVVGICDLARLG
PPEPMMAECLEKFTVSLNHKLDSHAELLDATQHTLQQQIQTLVKEGLRGFREARRDFWRGAESLEAALTH
NAEVPRRRAQEAEEAGAALRTARAGYRGRALDYALQINVIEDKRKFDIMEFVLRLVEAQATHFQQGHEEL
SRLSQYRKELGAQLHQLVLNSAREKRDMEQRHVLLKQKELGGEEPEPSLREGPGGLVMEGHLFKRASNAF
KTWSRRWFTIQSNQLVYQKKYKDPVTVVVDDLRLCTVKLCPDSERRFCFEVVSTSKSCLLQADSERLLQL
WVSAVQSSIASAFSQARLDDSPRGPGQGSGHLAIGSAATLGSGGMARGREPGGVGHVVAQVQSVDGNAQC
CDCREPAPEWASINLGVTLCIQCSGIHRSLGVHFSKVRSLTLDSWEPELVKLMCELGNVIINQIYEARVE
AMAVKKPGPSCSRQEKEAWIHAKYVEKKFLTKLPEIRGRRGGRGRPRGQPPVPPKPSIRPRPGSLRSKPE
PPSEDLGSLHPGALLFRASGHPPSLPTMADALAHGADVNWVNGGQDNATPLIQATAANSLLACEFLLQNG
ANVNQADSAGRGPLHHATILGHTGLACLFLKRGADLGARDSEGRDPLTIAMETANADIVTLLRLAKMREA
EAAQGQAGDETYLDIFRDFSLMASDDPEKLSRRSHDLHTL

Human CENTB1 protein sequence - var2 (public gi: 34533015) (SEQ ID NO: 232)
MSALAVSMAMVRGSLPSESRAPRSAPRFRNRQASLERRARVSRPPNFSQPSSPCHHPYPVWPRMVAWCSG
PRPALSAWFTFAPFWRRNSQARREFCMEKSRRGVEGGIPSGGFQDVLGWRQFREWEGGVW

Human CENTB1 pray sequence - var1 (SEQ ID NO: 41)
GCCTGGAGTACCCATACGACGTACCAGATTACGCTCATATGGCCATGGAGGCCAGTGAATTCCACCCAAG
CAGTGGTATCAACGCAGAGTGGCCATTATGGCCGGGGAAGGAGGCCTGGATTCACGCTAAATACGTGGAG
AAGAAGTTCCTGACCAAGCTGCCTGAGATTCGAGGGCGAAGAGGTGGCCGGGGGCGCCCAAGGGGGCAGC
CTCCTGTGCCCCCAAAGCCTTCCATCAGGCCCCGGCCAGGGAGCTTGAGATCCAAGCCAGAGCCCCCCTC
TGAGGACCTGGGAAGCCTGCACCCTGGGGCCCTACTGTTTCGAGCGTCTGGGCATCCTCCATCTCTTCCC
ACCATGTCGGCCGCCTCGGCCTCTAGAGGGTGGGCATCGATACGGGATCCATCGAGCTCGAGCTGCAGAT
GAATCGTAGATACTGAAAAACCCCGCAAGTTCACTTCAACTGTGCATTCGTGC

Human CENTB1 pray sequence - var2 (SEQ ID NO: 42)
CCGGCATGAGTACCATAACGACGTACAGATTACAGCGTNCATATAGTGACCATGGAGGCAGTGAATTCCA
CCGCAAGCAGTGGTATCAACGTCATGAGATGGACCATTATGAGCCGGGGTGGGCAGCCTCCTGATGTCCC
CGCGAAAGGCCTTCCATCAGGCNCCGGCAGAGGCAGCTTGAGATCCAAGCCAAGAGCCCCCCTCTGAGGA
CCTGGGTAAGACCTGCTACCACTAGTGCGGCCCTACTGTTNCGAGCGTCTGGGCATACTCCATCTCTTCC
CAACCGATGGNCTGATGCCCTTTGGCGCCATGGTAGCTTGATGTCAACCTAGGTGTACAANTGTGAGTGG
CCTNAAGGATAAATTGCTCGTCGACCAGACCGGCTATCCAAAGGCACAATAATCTAGCTAATTCGTTACG
TTCTTGG

Human CENTB1 pray sequence - var3 (SEQ ID NO: 43)
CCTGGAGTACCCATACGACGTACCAGATTACGCTCATATGGCCATGGAGGCCAGTGAATTCCACCCAAGC
AGTGGTATCAACGCAGAGTGGCCATTATGGCCGGGGGGGCAGCCTCCTGTGCCCCCAAAGCCTTCCATCA
GGCCCCGGCCAGGGAGCTTGAGATCCAAGCCAGAGCCCCCCTCTGAGGACCTGGGAAGCCTGCACCCTGG
GGCCCTACTGTTTCGAGCGTCTGGGCATCCTCCATCTCTTCCCACCATGGCTGATGCCCTTGCCCATGGA
GCTGATGTCAACTGGGTCAATGGGGGCCAAGATAATGCCACACCGCCGATCCAGGCCACAGCTGCTAATT
CTCTTCTGGCCTGTGAGTTTNNGCTCCAGAACGGGGCGAACGTGAACCAAGCGGACAGTGCGGGCCGGGG
CCCGCTGCACCACGCAACCATTCTTGGCCACACGGGGCTCGCCTGCCTGTTCCTGAAACGGGGAGCGGAT
CTGGGGGCTCGAGACTCTGAAGGCAGGGACCCTCTGACCATCGCCATGGAAACAGCCAACGCTGACATCG
TCACCCTGCTACGACTGGCAAAGATGAGGGAGGCTGAAGCGGCCCAGGGGCAGGCAGGAGATGAGACGTA
TCTTGACATCTTCCGCGACTTCTCCCTCATGGCGTCAGACNACCCNGAGAAGCTGANCCGTCGCAGTCAT
GACCTCCACACGCTGTGACCCGAGGCCCACGGGGCCCGCGCCTGCCTTCNTTTCCCGNCACCGGGCCCTT
TGNCATNAAAGCCTNCGNGCTTCNAAAAAAAAAAAAAAAAAAAA

Human CENTB1 pray sequence - var4 (SEQ ID NO: 44)
CCGGCCATGGAGTACCATACGACGTACAGTATTACAGCTACATATGGCCATGGAGGCCAGTGAATTCCAC
CGCAATGCAGTGGTATCAACGCATGCAGATGGACCATTATGGCCGGGGTGGGCACGTCCGTCCATGATGT
CCCCCAAAGGCCTTCCATCAGGCCCCGTGGCAGAGGAGGCTTGAGATCCAAGCCAGAGCCCACCCTCTGA
GGACCTGGGAAGCCTGCACCCNGGCGGCCCTACTGTTTCGAGCGTCTGGGACATACTCCATCATCTTCCC

ACGCGATGGACTGATGCCCTTGGGCCAATGGACGCTGATGTCAACTGGTGTACAGAGTGTGAGTGGCCAA
GATTAACTGCTCATCACCCGATGATCCATGGCCACTAGTCTGCTAAATATCTCTTCTGGCCTGTGAGTTT
CTCCTCACAGAAACGGTGCGCTGCAATCGTGAACNCAAAGCGGATCGAGTTGCAGGGCCTGGGGCCCGNG
TTGCACCGATCGCAAGCCAATTCTTGGCCANCTATCTGCGGGCTCGCCTGCCTGTTCCTGANACGAGGGA
GCTGATCTGGGGCGCTCGACGACTCTGAAG

Human CENTB1 pray sequence - var5 (SEQ ID NO: 45)
GCCATGGATACCATACGACGTACAGATTACGCTCATATGGCCATGGAGGCAGTGAATTCCACCCAAGCAG
TGGTATCAACGCATGAGATGGTCATTATGGCCGGGCAGGAGAAGGAGGCCTGGATTCACGCTAAATACG
TGGAGAAGAAGTTCCTGACCAAGCTGCCGTGAGAATTCGAGGGCGAAGAGGTGGCCGNGGGCGCCCAAGG
GGGCAGCCTCCTGTGCCAGCCCTAAAGCCTTCCATCATGGCCCCGCGTCCAAGGAGCTTGAGATCCAATG
CCGAGTAGCCCCCCTCTGACGGACCTAGGGAAGCCTGCTACCCTGAGGTGCCCTACGTGTTTCGAGCGTC
TGGGCATCCTCCATCTCTTTCCACCATGGCCTGATGCCCTTGCCCATGGAGCTGATGTCAACTGGGTCAA
TGGGTGGCCAAGATATATGCCACACCGCTGATCCAGTGCCACAGCTGCTACTTCTCTTCTGGCCTGTTGA
NTNTTCTCCTCCAGAACGGTGGCGACACGTGAACCCAAGCGGNCAGTGCCGGC

Human CENTB1 pray sequence - var6 (SEQ ID NO: 46)
GGCCATGGAGTACCATACGACGTACAGATTACGCTCATATGGCCATGGAGGCCAGTGAATTCCACCGCAA
GCAGTGGTATCAACGCATGAGATGGACCATTATGGGGGCAGTGCCATGGGCAGCTGAAGAAATCCANGC
CCAGCTGCTCCCGGCAGGAGAAGGAGGCCTGGATTCACGGCATAATAGTAGCAGCTGGAGTAAGAAGTTC
CTGTATCCAAGTCTGCCCTGACGAATTCGAGGTGGCGAAGTATGGTGGCCGGGGGCAGTCCTCGAAGGAG
GGTCAGCCACTCCTGGTGCCGCCACGAACATGCCCGTTTCCATACACGCGTCCCCGGCCCACGGGATGGC
ATTGAGATCCACATGCACAGAGCCCCGCCTCTGAGGACCTGTGAGCAAGCTCATGGCAACCCTGGGGACC
CTAGCGTAGTATTCTGAGCCAGTCTGGGCAATCGCTTCACATCTCTTCTCCACGCATGAGCATGATGCGC
GCTTTGACCCATGGAGCTAGATGTCAACTGGGTCAATGGGGTGCCAAGATAATCGCCACACCGTCTGATC
CAAGGCCTACAGCTGCTAACGTTCCTCTTCTGGCCTGTGAGTTTCTCCTCTCAGAACGGGGCGAACTGTG
AAGCCCAAGCGTGACAGTGCGGGCCCGGGGCCCGACTGCGCCACGCAATCCACTTCTTGGCCNGCAACNT
GGGCTCGNCTTGCCTTGTTTCCTTGATCAC

Human CENTB1 pray sequence - var7 (SEQ ID NO: 47)
CNCGGCATGGAGTACCATACGACGTACAGATTACGCTCATATGGCCATGGAGGCAGTGAATTCCACCCAA
GCAGTGGTATCAACGCATGAGTGGACCATTATGGGGAAGCTCATGTGTGAGCATGGGAAATAGTCATCA
TCAACCAAGATCTATGAGGCCCGCGTGGAGGCCATGGCAGTGAAGAAACCAGGGCCCAGTCTGCTCCCGG
CAGGAGAAGGAGGCCTGGATTCACGCTAAATACGTGGAAGAAGAAGTTCCTGACCAAGCTGCCTGAGATT
CGATGGCGANGAGGTGGCCGGGGGCGCCCAANGGGGCAGNCTCCTGTGCCCCCAAAGCCTTCCATCAGGC
CCCAGGCGCAGGGAGCTTGAGATCCAATGCCAGAGCCCCCCGTCTGAGGACCTGGGAAGCCTGCACCCTG
GGGCCCTACTGTTTCGAGCGTCTGGGCATCCTCCATCTCTTCCCACCATGGCTGATGCCCTTGCCCATGG
AGCTGATGTCAACTGGGTCAATGNGGCGGCCAAGATAATGCCATCACCGACTGATCCAGGCCACAGCCTG
CTAANTTCTACTTCTGGCCGTGTGAGTTTCTCCTCCAGGAACGGGGCGAACCGTGGACCAAGGCGGACNN
GTGCGGGCCGGGGCCCGCTGCCACCACGCCAACCATTCTTGGCATACGGGGCTCGCCT

Unigene Name: DDEF1 **Unigene ID:** Hs.386779

Human DDEF1 mRNA sequence - var1 (public gi: 31873727) (SEQ ID NO: 48)
GAGACAAAGTTTACAAAAATTGAGAAAGAGAAAAGAGAGCACGCAAAACAACATGGATGATCCGCACAG
AGATAACAGGAGCTGAGATTGCGGAAGAAATGGAGAAGGAAAGGCGCCTCTTTCAGCTCCAAATGTGTGA
ATATCTCATTAAAGTTAATGAAATCAAGACCAAAAAGGGTGTGGATCTGCTGCAGAATCTTATAAAGTAT
TACCATGCACAGTGCAATTTCTTTCAAGATGGCTTGAAAACAGCTGATAAGTTGAAACAGTACATTGAAA
AACTGGCTGCTGATTTATATAATATAAAACAGACCCAGGATGAAGAAAAGAAACAGCTAACTGCACTCCG
AGACTTAATAAAATCCTCTCTTCAACTGGATCAGAAAGAAGATTCTCAGAGCCGGCAGGGAGGATACAGC
ATGCATCAGCTCCAGGGCAATAAGGAATATGGCAGTGAAAAGAAGGGGTACCTGCTAAAGAAAAGTGACG
GGATCCGGAAAGTATGGCAGAGGAGGAAGTGTTCAGTCAAGAATGGGATTCTGACCATCTCACATGCCAC
ATCTAACAGGCAACCAGCCAAGTTGAACCTTCTCACCTGCCAAGTAAAACCTAATGCCGAAGACAAAAAA
TCTTTTGACCTGATATCACATAATAGAACATATCACTTTCAGGCAGAAGATGAGCAGGATTATGTAGCAT
GGATATCAGTATTGACAAATAGCAAAGAAGAGGCCCTAACCATGGCCTTCCGTGGAGAGCAGAGTGCGGG
AGAGAACAGCCTGGAAGACCTGACAAAAGCCATTATTGAGGATGTCCAGCGGCTCCCAGGGAATGACATT
TGCTGCGATTGTGGCTCATCAGAACCCACCTGGCTTTCAACCAACTTGGGTATTTTGACCTGTATAGAAT
GTTCTGGCATCCATAGGGAAATGGGGGTTCATATCTCTCGCATTCAGTCTTTGGAACTAGACAAATTAGG
AACTTCTGAACTCTTGCTGGCCAAGAATGTAGGAAACAATAGTTTTAATGATATTATGGAAGCAAATTTA
CCCAGCCCCTCACCAAAACCCACCCCTTCAAGTGATATGACTGTACGAAAAGAATATATCACTGCAAAGT
ATGTAGATCATAGGTTTTCAAGGAAGACCTGTTCAACTTCATCAGCTAAACTAAATGAATTGCTTGAGGC

CATCAAATCCAGGGATTTACTTGCACTAATTCAAGTCTATGCAGAAGGGGTAGAGCTAATGGAGCCACTG
CTGGAACCTGGGCAGGAGCTTGGGGAGACAGCCCTTCACCTTGCCGTCCGAACTGCAGATCAGACATCTC
TCCATTTGGTTGACTTCCTTGTACAAAACTGTGGGAACCTGGATAAGCAGACGGCCCTGGGAAACACAGT
TCTACACTACTGTAGTATGTACAGTAAACCTGAGTGTTTGAAGCTTTTGCTCAGGAGCAAGCCCACTGTG
GATATAGTTAACCAGGCTGGAGAAACTGCCCTAGACATAGCAAAGAGACTAAAAGCTACCCAGTGTGAAG
ATCTGCTTTCCCAGGCTAAATCTGGAAAGTTCAATCCACACGTCCACGTAGAATATGAGTGGAATCTTCG
ACAGGAGGAGATAGATGAGAGCGATGATGATCTGGATGACAAACCAAGCCCTATCAAGAAAGAGCGCTCA
CCCAGACCTCAGAGCTTCTGCCACTCCTCCAGCATCTCCCCCAGGACAAGCTGGCACTGCCAGGATTCA
GCACTCCAAGGGACAAACAGCGGCTCTCCTATGGAGCCTTCACCAACCAGATCTTCGTTTCCACAAGCAC
AGACTCGCCCACATCACCAACCACGGAGGCTCCCCCTCTGCCTCCTAGGAACGCCGGGAAAGGTCCAACT
GGCCCACCTTCAACACTCCCTCTAAGCACCCAGACCTCTAGTGGCAGCTCCACCCTATCCAAGAAGAGGC
CTCCTCCCCCACCACCCGGACACAAGAGAACCCTATCCGACCCTCCCAGCCCACTACCTCATGGGCCCCC
AAACAAAGGCGCAGTTCCTTGGGGTAACGATGGGGGTCCATCCTCTTCAAGTAAGACTACAAACAAGTTT
GAGGGACTATCCCAGCAGTCGAGCACCAGTTCTGCAAAGACTGCCCTTGGCCCAAGAGTTCTTCCTAAAC
TACCTCAGAAAGTGGCACTAAGGAAAACAGATCATCTCTCCCTAGACAAAGCCACCATCCCGCCCGAAAT
CTTTCAGAAATCATCACAGTTGGCAGAGTTGCCACAAAAGCCACCACCTGGAGACCTGCCCCCAAAGCCC
ACAGAACTGGCCCCCAAGCCCCAAATTGGAGATTTGCCGCCTAAGCCAGGAGAACTGCCCCCCAAACCAC
AGCTGGGGGACCTGCCACCCAAACCCCAACTCTCAGACTTGCCTCCCAAACCACAGATGAAGGACCTGCC
CCCCAAACCACAGCTGGGAGACCTGCTAGCAAAATCCCAGACTGGAGATGTCTCACCCAAGGCTCAGCAA
CCCTCTGAGGTCACACTGAAGTCACACCCATTGGATCTATCCCCAAATGTGCAGTCCAGAGACGCCATCC
AAAAGCAAGCATCTGAAGACTCCAACGACCTCACGCCTACTCTGCCAGAGACGCCCGTACCACTGCCCAG
AAAAATCAATACGGGGAAAAATAAAGTGAGGCGAGTGAAGACCATTTATGACTGCCAGGCAGACAACGAT
GACGAGCTCACATTCATCGAGGGAGAAGTGATTATCGTCACAGGGGAAGAGGACCAGGAGTGGTGGATTG
GCCACATCGAAGGACAGCCTGAAAGGAAGGGGGTCTTTCCAGTGTCCTTTGTTCATATCCTGTCTGACTA
GCAAAACGCAGAACCTTAAGATTGTCCACATCCTTCATGCAAGACTGCTGCCTTCATGTAACCCTGGGCA
CAGTGTGTATATAGCTGCTGTTACAGAGTAAGAAACTCATGGAAGGGCCACCTCAGGAGGGGGATATAAT
GTGTGTTGTAAATATCCTGTGGTTTTCTGCCTTCACCAGTATGAGGGTAGCCTCGGACCCGGCGCGCCTT
ACTGGTTTGCCAAAGCCATCCTTGGCATCTAGCACTTACATCTCTCTATGCTGTTCTACAAGCAAACAAA
CAAAAATAGGAGTATAGGAACTGCTGGCTTTGCAAATAGAAGTGGTCTCCAGCAACCGTTGAAAGGCATA
GAATTGACTCTGTTCCTAACAATGCAGTATTCTCAATTGTGTTACTGAAAATGCAACATTAGCAAAGAGG
TGGGTTCTGTTTTCCAGGTGAAACTTTTAGCTCCATGACAGACCAGCCTGTAGTTATCTGTGTACACAGT
TTACAGCTACAAAAACCTACTTTGGTATTTATTACAGAAAAGTGCTCAGTTAAATGTAAGTGTTATTCCT
TCAGCAAAATATTCACTGACCCAAAACTCTTTATGGCATTTTACAATGCACACAGCCTCATGCAAGTTTA
GACAAGTGGATTTATACTGTCTTATGAGTGCCCGCCCCTGATATATTACCTCATTATGCAAAAATAACAT
ATCTTTCATGACTATTTTGACAAAAGTTTAAAACACATATGAAGTTCAAATTTCAGGAACCAAGGACTGC
CAGAAAATATTAGCCTCTACATTACGCATGCATTTAGAAACTTACCTGAAATCTGCCTTTTATAAAGGAA
TAGTATGGATAAGTGGAATTGTACATTTTTTAAACTTGATTGCCATTAAAGCAGAAATTATAAGGTTGCA
ACAATATTTGTTTCTAATCACTGGCTTTCTCAAGAGTATGGATTGACATATTGTGTTATGAATGCATATC
TCTCAGATGTGTTGAAGCATCCATTGCATCCATTTTTTATTATTTTCTTAGTTTTGTTCTTGGACAAATT
TAAACTTTTAAAAGATTATTCAAGATGAATTTAAAAGTCAACCCTTCACACAGTTTCCCTACTGTATGTA
GAATCCAGGTGCTGAAACCAAGTGTTTCTTTTCCCATGCTCTTTGTTAAACCCCAATTATAGATAATTTT
TCCAGTCTTAAGCTCTGTCCACCTTCAAGTCAATTCATAACCAAGTTTTTGAACGCTGCTATGAATTGCA
CTGTGAAAAGCACTCTTCCCTCTCAGTTTTCTTTTCATCCCAGCCATGTTTATCAGATCCTTAAGAACAT
TGTATTTCAGTCTTTTACATCAGTCTGAATTTTGGAAAAGAATGCAATAGTTGTACTCCACAGTCAGTGG
AACTGTTCCCTGAGTCCGAGGCTCATGTGTCATTCTGGCACTACATTTGCTTAAATTGCTATTTTGGCAA
CAGCACAGAAAACTAATATTTTTAAGCAGAGAATCTTGGCAATGAGTGAGAGATGTTAATTTCACAGAAG
CACAACTCCCAACCCAACCCTTAGGAAAAGCCCTCTTCCATCGTTACAGTGCTCAGTGAATATTAATTTA
GTTCTGCTTAAGTGGTTGCTATACAAACTTTGAATAGCCACCTAATAAATAAACCTTGCATGACAAACCT
GCAAAATATTTTATCAGCTGTTATTGGAAAGTGATTTTAAGCAATTGCTTCCTCAGTGTCAGGGCACATG
TGAATTTCCACACCAAACAGAGCATGAGGAACCAGTTGACATGCTGGGTTGTGACTGGCAGCTTTAGCAG
CCTCGGTACTGAAGCCACACCAGTGTCCGGATGGAAGTCTGCATCTGAGGTTGCTCAGTGTCCCGGTCAT
TCATTTACACATTTTAACTTGCATTAAAGAGCTGTTCTTTTCTGTGGCCTAGACTCTTTTCACTGATCTC
AAAATAAACTGGTTTTTTTCAAAAAAAAAAAAAAAAAAA

Human DDEF1 mRNA sequence - var2 (public gi: 6330853) (SEQ ID NO: 49)
GAAAAGAGAGCACGCAAAACAACATGGGATGATCCGCACAGAGATAACAGGAGCTGAGATTGCGGAAGAA
ATGGAGAAGGAAAGGCGCCTCTTTCAGCTCCAAATGTGTGAATATCTCATTAAAGTTAATGAAATCAAGA
CCAAAAAGGGTGTGGATCTGCTGCAGAATCTTATAAAGTATTACCATGCACAGTGCAATTTCTTTCAAGA
TGGCTTGAAAACAGCTGATAAGTTGAAACAGTACATTGAAAAACTGGCTGCTGATTTATATAATATAAAA
CAGACCCAGGATGAAGAAAAGAAACAGCTAACTGCACTCCGAGACTTAATAAAATCCTCTCTTCAACTGG
ATCAGAAAGAATCTAGGAGAGATTCTCAGAGCCGGCAAGGAGGATACAGCATGCATCAGCTCCAGGGCAA
TAAGGAATATGGCAGTGAAAAGAAGGGGTACCTGCTAAAGAAAAGTGACGGGATCCGGAAAGTATGGCAG
AGGAGGAAGTGTTCAGTCAAGAATGGGATTCTAACCATCTCACATGCCACATCTAACAGGCAACCAGCCA

AGTTGAACCTTCTCACCTGCCAAGTAAAACCTAATGCCGAAGACAAAAAATCTTTTGACCTGATATCACA
TAATAGAACATATCACTTTCAGGCAGAAGATGAGCAGGATTATGTAGCATGGATATCAGTATTGACAAAT
AGCAAAGAAGAGGCCCTAACCATGGCCTTCCGTGGAGAGCAGAGTGCGGGAGAGAACAGCCTGGAAGACC
TGACAAAAGCCATTATTGAGGATGTCCAGCGGCTCCCAGGGAATGACATTTGCTGCGATTGTGGCTCATC
AGAACCCACCTGGCTTTCAACCAACTTGGGTATTTTGACCTGTATAGAATGTTCTGGCATCCATAGGGAA
ATGGGGGTTCATATTTCTCGCATTCAGTCTTTGGAACTAGACAAATTAGGAACTTCTGAACTCTTGCTGG
CCAAGAATGTAGGAAACAATAGTTTTAATGATATTATGGAAGCAAATTTACCCAGCCCCTCACCAAAACC
CACCCCTTCAAGTGATATGACTGTACGAAAAGAATATATCACTGCAAAGTATGTAGATCATAGGTTTTCA
AGGAAGACCTGTTCAACTTCATCAGCTAAACTAAATGAATTGCTTGAGGCCATCAAATCCAGGGATTTAC
TTGCACTAATTCAAGTCTATGCAGAAGGGGTAGAGCTAATGGAACCACTGCTGGAACCTGGGCAGGAGCT
TGGGGAGACAGCCCTTCACCTTGCCGTCCGAACTGCAGATCAGACATCTCTCCATTTGGTTGACTTCCTT
GTACAAAACTGTGGGAACCTGGATAAGCAGACGGCCCTGGGAAACACAGTTCTACACTACTGTAGTATGT
ACAGTAAACCTGAGTGTTTGAAGCTTTTGCTCAGGAGCAAGCCCACTGTGGATATAGTTAACCAGGCTGG
AGAAACTGCCCTAGACATAGCAAAGAGACTAAAAGCTACCCAGTGTGAAGATCTGCTTTCCCAGGCTAAA
TCTGGAAAGTTCAATCCACACGTCCACGTAGAATATGAGTGGAATCTTCGACAGGAGGAGATAGATGAGA
GCGATGATGATCTGGATGACAAACCAAGCCCTATCAAGAAAGAGCGCTCACCCAGACCTCAGAGCTTCTG
CCACTCCTCCAGCATCTCCCCCCAGGACAAGCTGGCACTGCCAGGATTCAGCACTCCAAGGGACAAACAG
CGGCTCTCCTATGGAGCCTTCACCAACCAGATCTTCGTTTCCACAAGCACAGACTCGCCCACATCACCAA
CCACGGAGGCTCCCCCTCTGCCTCCTAGGAACGCCGGGAAAGGTCCAACTGGCCCACCTTCAACACTCCC
TCTAAGCACCCAGACCTCTAGTGGCAGCTCCACCCTATCCAAGAAGAGGCCTCCTCCCCCACCACCCGGA
CACAAGAGAACCCTATCCGACCCTCCCAGCCCACTACCTCATGGGCCCCCAAACAAAGGCGCAGTTCCTT
GGGGTAACGATGGGGGTCCATCCTCTTCAAGTAAGACTACAAACAAGTTTGAGGGACTATCCCAGCAGTC
GAGCACCAGTTCTGCAAAGACTGCCCTTGGCCCAAGAGTTCTTCCTAAACTACCTCAGAAAGTGGCACTA
AGGAAAACAGATCATCTCTCCCTAGACAAAGCCACCATCCCGCCCGAAATCTTTCAGAAATCATCACAGT
TGGCAGAGTTGCCACAAAAGCCACCACCTGGAGACCTGCCCCCAAAGCCCACAGAACTGGCCCCCAAGCC
CCAAATTGGAGATTTGCCGCCTAAGCCAGGAGAACTGCCCCCCAAACCACAGCTGGGGGACCTGCCACCC
AAACCCCAACTCTCAGACTTACCTCCCAAACCACAGATGAAGGACCTGCCCCCCAAACCACAGCTGGGAG
ACCTGCTAGCAAAATCCCAGACTGGAGATGTCTCACCCAAGGCTCAGCAACCCTCTGAGGTCACACTGAA
GTCACACCCATTGGATCTATCCCCAAATGTGCAGTCCAGAGACGCCATCCAAAAGCAAGCATCTGAAGAC
TCCAACGACCTCACGCCTACTCTGCCAGAGACGCCCGTACCACTGCCCAGAAAAATCAATACGGGGAAAA
ATAAAGTGAGGCGAGTGAAGACCATTTATGACTGCCAGGCAGACAACGATGACGAGCTCACATTCATCGA
GGGAGAAGTGATTATCGTCACAGGGGAAGAGGACCAGGAGTGGTGGATTGGCCACATCGAAGGACAGCCT
GAAAGGAAGGGGGTCTTTCCAGTGTCCTTTGTTCATATCCTGTCTGACTAGCAAAACGCAGAACCTTAAG
ATTGTCCACATCCTTCATGCAAGACTGCTGCCTTCATGTAACCCTGGGCACAGTGTGTATATAGCTGCTG
TTACAGAGTAAGAAACTCATGGAAGGGCCACCTCAGGAGGGGGATATAATGTGTGTTGTAAATATCCTGT
GGTTTTCTGCCTTCACCAGTATGAGGGTAGCCTCGGACCCGGCGCGCCTTACTGGTTTGCCAAAGCCATC
CTTGGCATCTAGCACTTACATCTCTCTATGCTGTTCTACAAGCAAACAAACAAAAATAGGAGTATAGGAA
CTGCTGGCTTTGCAAATAGAAGTGGTCTCCAGCAACCGTTGAAAGGCATAGAATTGACTCTGTTCCTAAC
AATGCAGTATTCTCAATTGTGTTACTGAAAATGCAACATTAGCAAAGAGGTGGGTTCTGTTTTCCAGGTG
AAACTTTTAGCTCCATGACAGACCAGCCTGTAGTTATCTGTGTACACAGTTTACAGCTACAAAAACCTAC
TTTGGTATTTATTACAGAAAAGTGCTCAGTTAAATGTAAGTGTTATTCCTTCAGCAAAATATTCACTGAC
CCAAAACTCTTTATGGCATTTTACAATGCACACAGCCTCATGCAAGTTTAGACAAGTGGATTTATACTGT
CTTATGAGTGCCCGCCCCTGATATATTACCTCATTATGCAAAAATAACATATCTTTCATGACTATTTTGA
CAAAAGTTTAAAACACATATGAAGTTCAAATTTCAGGAACCAAGGACTGCCAGAAAATATTAGCCTCTAC
ATTACGCATGCATTTAGAAGCTTACCTGAAATCTGCCTTTTATAAAGGAATAGTATGGATAAGTGGAATT
GTACATTTTTTAAACTTGATTGCCATTAAAGCAGAAATTATAAGGTTGCAACAATATTTGTTTCTAATCA
CTGGCTTTCTCAAGAGTATGGATTGACATATTGTGTTATGAATGCACATCTCTCAGATGTGTTGAAGCAT
CCATTGCATCCATTTTTTATTATTTTCTTAGTTTTGTTCTTGGACAAATTTAAACTTTTAAAAGATTATT
CAAGATGAATTTAAAAGTCAACCCTTCACACAGTTTCCCTACTGTATGTAGAATCCAGGTGCTGAAACCA
AGTGTTTCTTTTCCCATGCTCTTTGTTAAACCCCAATTATAGATAATTTTTCCAGTCTTAAGCTCTGTCC
ACCTTCAAGTCAATTCATAACCAAGTTTTTGAACGCTGCTATGAATTGCACTGTGAAAAGCACTCTTCCC
TCTCAGTTTTCTTTTCATCCCAGCCATGTTTATCAGATCCTTAAGAACATTGTATTTCAGTCTTTTACAT
CAGTCTGAATTTTGGAAAAGAATGCAATAGTTGTACTCCACAGTCAGTGGAACTGTTCCCTGAGTCCGAG
GCTCATGTGTCATTCTGGCACTACATTTGCTTAAATTGCTATTTTGGCAACAGCACAGAAAACTAATATT
TTTAAGCAGAGAATCTTGGCAATGAGTGAGAGATGTTAATTTCACAGAAGCACAACTCCCAACCCAACCC
TTAGGAAAAGCCCTCTTCCATCGTTACAGTGCTCAGTGAATATTAATTTAGTTCTGCTTAAGTGGTTGCT
ATACAAACTTTGAATAGCCACCTAATAAATAAACCTTGCATGACAAACCTGCAAAATATTTTATCAGCTG
TTATTGGAAAGTGATTTTAAGCAATTGCTTCCTCAGTGTCAGGGCACATGTGAATTTCCACACCAAACAG
AGCATGAGGAACCAGTTGACATGCTGGGTTGTGACTGGCAGCTTTAGCAGCCTCGGTACTGAAGCCACAC
CAGTGTCCGGATGGAAGTCTGCATCTGAGGTTGCTCAGTGTCCCGGTCATTCATTTACACATTTTAACTT
GCATTAAAGAGCTGTTCTTTTCTGTGGCCTAGACTCTTTTCACTGATCTCAAAATAAACTGGTTTTTTTC
AAAAAAAAAAAAAAAACAAAAACAAAAAAAAAAACACAAAAGCTGCATGTCTAAAATTACATGGAGTTAGT
GTCTATTCTTTTTCCCCTTTTGCAGCAACTTACACAGCATTTTTAACACCTTTTTTTTCTAGTTTTTTTG

Figure 36 part - 25

TTCGGTTTTGTTTTCCATCAGGAATTTGAGTTCTCTCTAACCCAGCTTACTGTGGGACATAGGAAAACTC
AGTAGAAATACCTTTGGTGATCTTGTTGAGTTTAAGTCTGATCTTGATCTTAAACTCAGTAAGCCACTAT
CTGCAATTTTGTACATTATATAGTATTTTGAAGATATGGAACCTTATGAAAAAAAAATAGCAAATTAGTT
CTTTTTCCCCCAGAGGGGAAAGTTATGTTCTGCAAATAGTGTGTGTCTTATTTTACTGTTGAACAGCAAT
TGCTATTTATTTTTTTATTGCCTAGAACTTCAACATGTTGTATAGGAATCCTGTAGTGCCACTAGTTAAA
TGCCGAATTCTCATCTGGATGTTACCATCAAACATCAGTACACTTGTCATTTCACATGTGTTTAATGTGA
CAGTTTTTCAGTACTGTATGTGTTAATTTCTACTTTTTTTAATATTTAAAATTGCTTTTAAATAAACATA
TTCTCAGTTGATCCC

Human DDEF1 mRNA sequence - var3 (public gi: 7689053) (SEQ ID NO: 50)
GATTGCCATTAAAGCAGAAATTATAAGGTTGCAACAATATTTGTTTCTAATCACTGGCTTTCTCAAGAGT
ATGGATTGACATATTGTGTTATGAATGCACATCTCTCAGATGTGTTGAAGCATCCATTGCATCCATTTTT
TATTATTTTCTTAGTTTTGTTCTTGGACAAATTTAAACANNTTAAAAGATTATTCAAGATGAATTTAAAA
GTCAACCCTTCACACAGTTTCCCTACTGTATGTAGAATCCAGGTGCTGAAACCAAGTGTTTCTTTTCCCA
TGCTCTTTGTTAAACTCCAATTATAGATAATTTTTCCAGTCTTAAGCTCTGTCCACCTTCAAGTCAATTC
ATAACCAAGTTTTTGAACGCTGCTATGAATTGCACTGTGAAAAGCACTCTTCCCTCTCAGTTTTCGTTCA
TCCTGAGCCAGAATCAAAAAAAAAAA

Human DDEF1 mRNA sequence - var4 (public gi: 16552319) (SEQ ID NO: 51)
CAGAACCTTAAGATTGTCCACATCCTTCATGCAAGACTGCTGCCTTCATGTAACCCTGGGCACAGTGTGT
ATATAGCTGCTGTTACAGAGTAAGAAACTCATGGAAGGGCCACCTCAGGAGGGGGATATAATGTGTGTTG
TAAATATCCTGTGGTTTTCTGCCTTCACCAGTATGAGGGTAGCCTCGGACCCGGCGCGCCTTACTGGTTT
GCCAAAGCCATCCTTGGCATCTAGCACTTACATCTCTCTATGCTGTTCTACAAGCAAACAAACAAAAATA
GGAGTATAGGAACTGCTGGCTTTGCAAATAGAAGTGGTCTCCAGCAACCGTTGAAAGGCATAGAATTGAC
TCTGTTCCTAACAATGCAGTATTCTCAATTGTGTTACTGAAAATGCAACATTAGCAAAGAGGTGGGTTCT
GTTTTCCAGGTGAAACTTTTAGCTCCATGACAGACCAGCCTGTAGTTATCTGTGTACACAGTTTACAGCT
ACAAAAACCTACTTTGGTATTTATTACAGAAAAGTGCTCAGTTAAATGTAAGTGCTATTCCTTCAGCAAA
ATATTCACTGACCCAAAACTCTTTATGGCATTTTACAATGCACACAGCCTCATGCAAGTTTAGACAAGTG
GATTTATACTGTCTTATGAGTGCCCGCCCCTGATATATTACCTCATTATGCAAAAATAACATATCTTTCA
TGACTATTTTGACAAAAGTTTAAAACACATATGAAGTTCAAATTTCAGGAACCAAGGACTGCCAGAAAAT
ATTAGCCTCTACATTACGCATGCATTTAGAAGCTTACCTGAAATCTGCCTTTTATAAAGGAATAGTATGG
ATAAGTGGAATTGTACATTTTTTAAACTTGATTGCCATTAAAGCAGAAATTATAAGGTTGCAACAATATT
TGTTTCTAATCACTGGCTTTCTCAAGAGTATGGATTGACATATTGTGTTATGAATGCACATCTCTCAGAT
GTGTTGAAGCATCCATTGCATCCATTTTTTATTATTTTCTTAGTTTTGTTCTTGGACAAATTTAAACTTT
TAAAAGATTATTCAAGATGAATTTAAAAGTCAACCCTTCACACAGTTTCCCTACTGTATGTAGAATCCAG
GTGCTGAAACCAAGTGTTTCTTTTCCCATGCTCTTTGTTAAACCCCAATTATAGATAATTTTTCCAGTCT
TAAGCTCTGTCCACCTTCAAGTCAATTCATAACCAAGTTTTTGAACGCTGCTATGAATTGCACTGTGAAA
AGCACTCTTCCCTCTCAGTTTTCTTTTCATCCCAGCCATGTTTATCAGATCCTTAAGAACATTGTATTTC
AGTCTTTTACATCAGTCTGAATTTTGGAAAAGAATGCAATAGTTGTACTCCACAGTCAGTGGAACTGTTC
CCTGAGTCCGAGGCTCATGTGTCATTCTGGCACTACATTTGCTTAAATTGCTATTTTGGCAACAGCACAG
AAAACTAATATTTTTAAGCAGAGAATCTTGGCAATGAGTGAGAGATGTTAATTTCACAGAAGCACAACTC
CCAACCCAACCCTTAGGAAAAGCCCTCTTCCATCGTTACAGTGCTCAGTGAATATTAATTTAGTTCTGCT
TAAGTGGTTGCTATACAAACTTTGAATAGCCACCTAATAAATAAACCTTGCATGACAAACCTGCAAAATA
TTTTATCAGCTGTTATTGGAAAGTGATTTTAAGCAATTGCTTCCTCAGTGTCAGGGCACATGTGAATTTC
CACACCAAACAGAGCATGAGGAACCAGTTGACATGCTGGGTTGTGACTGGCAGCTTTAGCAGCCTCGGTA
CTGAAGCCACACCAGTGTCCGGATGGAAGTCTGCATCTGAGGTTGCTCAGTGTCCCGGTCATTCATTTAC
ACATTTTAACTTGCATTAAAGAGCTGTTCTTTTCTGTGGCCTAGACTCTTTTCACTGATCTCAAAATAAA
CTGGTTTTTTC

Human DDEF1 mRNA sequence - var5 (public gi: 18088817) (SEQ ID NO: 52)
CAGCTACAAAAACCTACTTTGGTATTTATTACAGAAAAGTGCTCAGTTAAATGTAAGTGTTATTCCTTCA
GCAAAATATTCACTGACCCAAAACTCTTTATGGCATTTTACAATGCACACAGCCTCATGCAAGTTTAGAC
AAGTGGATTTATACTGTCTTATGAGTGCCCGCCCCTGATATATTACCTCATTATGCAAAAATAACATATC
TTTCATGACTATTTTGACAAAAGTTTAAAACACATATGAAGTTCAAATTTCAGGAACCAAGGACTGCCAG
AAAATATTAGCCTCTACATTACGCATGCATTTAGAAGCTTACCTGAAATCTGCCTTTTATAAAGGAATAG
TATGGATAAGTGGAATTGTACATTTTTTAAACTTGATTGCCATTAAAGCAGAAATTATAAGGTTGCAACA
ATATTTGTTTCTAATCACTGGCTTTCTCAAGAGTATGGATTGACATATTGTGTTATGAATGCACATCTCT
CAGATGTGTTGAAGCATCCATTGCATCCATTTTTTATTATTTTCTTAGTTTTGTTCTTGGACAAATTTAA
ACTTTTAAAAGATTATTCAAGATGAATTTAAAAGTCAACCCTTCACACAGTTTCCCTACTGTATGTAGAA
TCCAGGTGCTGAAACCAAGTGTTTCTTTTCCCATGCTCTTTGTTAAACCCCAATTATAGATAATTTTTCC
AGTCTTAAGCTCTGTCCACCTTCAAGTCAATTCATAACCAAGTTTTTGAACGCTGCTATGAATTGCACTG
TGAAAAGCACTCTTCCCTCTCAGTTTTCTTTTCATCCCAGCCATGTTTATCAGATCCTTAAGAACATTGT

ATTTCAGTCTTTTACATCAGTCTGAATTTTGGAAAAGAATGCAATAGTTGTACTCCACAGTCAGTGGAAC
TGTTCCCTGAGTCCGAGGCTCATGTGTCATTCTGGCACTACATTTGCTTAAATTGCTATTTTGGCAACAG
CACAGAAAACTAATATTTTTAAGCAGAGAATCTTGGCAATGAGTGAGAGATGTTAATTTCACAGAAGCAC
AACTCCCAACCCAACCCTTAGGAAAAGCCCTCTTCCATCGTTACAGTGCTCAGTGAATATTAATTTAGTT
CTGCTTAAGTGGTTGCTATACAAACTTTGAATAGCCACCTAATAAATAAACCTTGCATGACAAACCTGCA
AAATATTTTATCAGCTGTTATTGGAAAGTGATTTTAAGCAATTGCTTCCTCAGTGTCAGGGCACATGTGA
ATTTCCACACCAAACAGAGCATGAGGAACCAGTTGACATGCTGGGTTGTGACTGGCAGCTTTAGCAGCCT
CGGTACTGAAGCCACACCAGTGTCCGGATGGAAGTCTGCATCTGAGGTTGCTCAGTGTCCCGGTCATTCA
TTTACACATTTTAACTTGCATTAAAGAGCTGTTCTTTTCTGTGGCCTAGACTCTTTTCACTGATCTCAAA
ATAAACTGGTTTTTTACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human DDEF1 mRNA sequence - var6 (Predicted by Proteologics) (SEQ ID NO: 53)

TTTCGACGCGTGGGTTTTATTCCCTTGAAGACTTTGGAAAGATTTTGTCATTCATCGCAATGATTGGTCA
GCCTCAAGAAGCATGCAGGAGCCATCATAAGAGTCACAAGGCTCTAGACCAAGATAGAACAGCCCTTCAG
AAAGTGAAGAAGTCTGTAAAAGCAATATATAATTCTGGTCAAGATCATGTACAAAATGAAGAAAACTATG
CACAAGTTCTTGATAAGTTTGGGAGTAATTTTTTAAGTCGAGACAACCCCGACCTTGGCACCGCGTTTGT
CAAGTTTTCTACTCTTACAAAGGAACTGTCCACACTGCTGAAAAATCTGCTCCAGGGTTTGAGCCACAAT
GTGATCTTCACCTTGGATTCTTTGTTAAAAGGAGACCTAAAGGGAGTCAAAGGAGATCTCAAGAAGCCAT
TTGACAAAGCCTGGAAAGATTATGAGACAAAGTTTACAAAAATTGAGAAAGAGAAAAGAGAGCACGCAAA
ACAACATGGGATGATCCGCACAGAGATAACAGGAGCTGAGATTGCGGAAGAAATGGAGAAGGAAAGGCGC
CTCTTTCAGCTCCAAATGTGTGAATATCTCATTAAAGTTAATGAAATCAAGACCAAAAAGGGTGTGGATC
TGCTGCAGAATCTTATAAAGTATTACCATGCACAGTGCAATTTCTTTCAAGATGGCTTGAAAACAGCTGA
TAAGTTGAAACAGTACATTGAAAAACTGGCTGCTGATTTATATAATATAAAACAGACCCAGGATGAAGAA
AAGAAACAGCTAACTGCACTCCGAGACTTAATAAAATCCTCTCTTCAACTGGATCAGAAAGAATCTAGGA
GAGATTCTCAGAGCCGGCAAGGAGGATACAGCATGCATCAGCTCCAGGGCAATAAGGAATATGGCAGTGA
AAAGAAGGGGTACCTGCTAAAGAAAAGTGACGGGATCCGGAAAGTATGGCAGAGGAGGAAGTGTTCAGTC
AAGAATGGGATTCTAACCATCTCACATGCCACATCTAACAGGCAACCAGCCAAGTTGAACCTTCTCACCT
GCCAAGTAAAACCTAATGCCGAAGACAAAAAATCTTTTGACCTGATATCACATAATAGAACATATCACTT
TCAGGCAGAAGATGAGCAGGATTATGTAGCATGGATATCAGTATTGACAAATAGCAAAGAAGAGGCCCTA
ACCATGGCCTTCCGTGGAGAGCAGAGTGCGGGAGAGAACAGCCTGGAAGACCTGACAAAAGCCATTATTG
AGGATGTCCAGCGGCTCCCAGGGAATGACATTTGCTGCGATTGTGGCTCATCAGAACCCACCTGGCTTTC
AACCAACTTGGGTATTTTGACCTGTATAGAATGTTCTGGCATCCATAGGGAAATGGGGGTTCATATTTCT
CGCATTCAGTCTTTGGAACTAGACAAATTAGGAACTTCTGAACTCTTGCTGGCCAAGAATGTAGGAAACA
ATAGTTTTAATGATATTATGGAAGCAAATTTACCCAGCCCCTCACCAAAACCCACCCCTTCAAGTGATAT
GACTGTACGAAAAGAATATATCACTGCAAAGTATGTAGATCATAGGTTTTCAAGGAAGACCTGTTCAACT
TCATCAGCTAAACTAAATGAATTGCTTGAGGCCATCAAATCCAGGGATTTACTTGCACTAATTCAAGTCT
ATGCAGAAGGGGTAGAGCTAATGGAACCACTGCTGGAACCTGGGCAGGAGCTTGGGGAGACAGCCCTTCA
CCTTGCCGTCCGAACTGCAGATCAGACATCTCTCCATTTGGTTGACTTCCTTGTACAAAACTGTGGGAAC
CTGGATAAGCAGACGGCCCTGGGAAACACAGTTCTACACTACTGTAGTATGTACAGTAAACCTGAGTGTT
TGAAGCTTTTGCTCAGGAGCAAGCCCACTGTGGATATAGTTAACCAGGCTGGAGAAACTGCCCTAGACAT
AGCAAAGAGACTAAAAGCTACCCAGTGTGAAGATCTGCTTTCCCAGGCTAAATCTGGAAAGTTCAATCCA
CACGTCCACGTAGAATATGAGTGGAATCTTCGACAGGAGGAGATAGATGAGAGCGATGATGATCTGGATG
ACAAACCAAGCCCTATCAAGAAAGAGCGCTCACCCAGACCTCAGAGCTTCTGCCACTCCTCCAGCATCTC
CCCCCAGGACAAGCTGGCACTGCCAGGATTCAGCACTCCAAGGGACAAACAGCGGCTCTCCTATGGAGCC
TTCACCAACCAGATCTTCGTTTCCACAAGCACAGACTCGCCCACATCACCAACCACGGAGGCTCCCCCTC
TGCCTCCTAGGAACGCCGGGAAAGGTCCAACTGGCCCACCTTCAACACTCCCTCTAAGCACCCAGACCTC
TAGTGGCAGCTCCACCCTATCCAAGAAGAGGCCTCCTCCCCCACCACCCGGACACAAGAGAACCCTATCC
GACCCTCCCAGCCCACTACCTCATGGGCCCCCAAACAAAGGCGCAGTTCCTTGGGGTAACGATGGGGGTC
CATCCTCTTCAAGTAAGACTACAAACAAGTTTGAGGGACTATCCCAGCAGTCGAGCACCAGTTCTGCAAA
GACTGCCCTTGGCCCAAGAGTTCTTCCTAAACTACCTCAGAAAGTGGCACTAAGGAAAACAGATCATCTC
TCCCTAGACAAAGCCACCATCCCGCCCGAAATCTTTCAGAAATCATCACAGTTGGCAGAGTTGCCACAAA
AGCCACCACCTGGAGACCTGCCCCCAAAGCCCACAGAACTGGCCCCCAAGCCCCAAATTGGAGATTTGCC
GCCTAAGCCAGGAGAACTGCCCCCCAAACCACAGCTGGGGGACCTGCCACCCAAACCCCAACTCTCAGAC
TTACCTCCCAAACCACAGATGAAGGACCTGCCCCCCAAACCACAGCTGGGAGACCTGCTAGCAAAATCCC
AGACTGGAGATGTCTCACCCAAGGCTCAGCAACCCTCTGAGGTCACACTGAAGTCACACCCATTGGATCT
ATCCCCAAATGTGCAGTCCAGAGACGCCATCCAAAAGCAAGCATCTGAAGACTCCAACGACCTCACGCCT
ACTCTGCCAGAGACGCCCGTACCACTGCCCAGAAAAATCAATACGGGGAAAAATAAAGTGAGGCGAGTGA
AGACCATTTATGACTGCCAGGCAGACAACGATGACGAGCTCACATTCATCGAGGGAGAAGTGATTATCGT
CACAGGGGAAGAGGACCAGGAGTGGTGGATTGGCCACATCGAAGGACAGCCTGAAAGGAAGGGGGTCTTT
CCAGTGTCCTTTGTTCATATCCTGTCTGACTAGCAAAACGCAGAACCTTAAGATTGTCCACATCCTTCAT
GCAAGACTGCTGCCTTCATGTAACCCTGGGCACAGTGTGTATATAGCTGCTGTTACAGAGTAAGAAACTC
ATGGAAGGGCCACCTCAGGAGGGGATATAATGTGTGTTGTAAATATCCTGTGGTTTTCTGCCTTCACCA
GTATGAGGGTAGCCTCGGACCCGGCGCGCCTTACTGGTTTGCCAAAGCCATCCTTGGCATCTAGCACTTA

CATCTCTCTATGCTGTTCTACAAGCAAACAAACAAAAATAGGAGTATAGGAACTGCTGGCTTTGCAAATA
GAAGTGGTCTCCAGCAACCGTTGAAAGGCATAGAATTGACTCTGTTCCTAACAATGCAGTATTCTCAATT
GTGTTACTGAAAATGCAACATTAGCAAAGAGGTGGGTTCTGTTTTCCAGGTGAAACTTTTAGCTCCATGA
CAGACCAGCCTGTAGTTATCTGTGTACACAGTTTACAGCTACAAAAACCTACTTTGGTATTTATTACAGA
AAAGTGCTCAGTTAAATGTAAGTGTTATTCCTTCAGCAAAATATTCACTGACCCAAAACTCTTTATGGCA
TTTTACAATGCACACAGCCTCATGCAAGTTTAGACAAGTGGATTTATACTGTCTTATGAGTGCCCCGCCCC
TGATATATTACCTCATTATGCAAAAATAACATATCTTTCATGACTATTTTGACAAAAGTTTAAAACACAT
ATGAAGTTCAAATTTCAGGAACCAAGGACTGCCAGAAAATATTAGCCTCTACATTACGCATGCATTTAGA
AGCTTACCTGAAATCTGCCTTTTATAAAGGAATAGTATGGATAAGTGGAATTGTACATTTTTTAAACTTG
ATTGCCATTAAAGCAGAAATTATAAGGTTGCAACAATATTTGTTTCTAATCACTGGCTTTCTCAAGAGTA
TGGATTGACATATTGTGTTATGAATGCACATCTCTCAGATGTGTTGAAGCATCCATTGCATCCATTTTTT
ATTATTTTCTTAGTTTTGTTCTTGGACAAATTTAAACTTTTAAAAGATTATTCAAGATGAATTTAAAAGT
CAACCCTTCACACAGTTTCCCTACTGTATGTAGAATCCAGGTGCTGAAACCAAGTGTTTCTTTTCCCATG
CTCTTTGTTAAACCCCAATTATAGATAATTTTTCCAGTCTTAAGCTCTGTCCACCTTCAAGTCAATTCAT
AACCAAGTTTTTGAACGCTGCTATGAATTGCACTGTGAAAAGCACTCTTCCCTCTCAGTTTTCTTTTCAT
CCCAGCCATGTTTATCAGATCCTTAAGAACATTGTATTTCAGTCTTTTACATCAGTCTGAATTTTGGAAA
AGAATGCAATAGTTGTACTCCACAGTCAGTGGAACTGTTCCCTGAGTCCGAGGCTCATGTGTCATTCTGG
CACTACATTTGCTTAAATTGCTATTTTGGCAACAGCACAGAAAACTAATATTTTTAAGCAGAGAATCTTG
GCAATGAGTGAGAGATGTTAATTTCACAGAAGCACAACTCCCAACCCAACCCTTAGGAAAAGCCCTCTTC
CATCGTTACAGTGCTCAGTGAATATTAATTTAGTTCTGCTTAAGTGGTTGCTATACAAACTTTGAATAGC
CACCTAATAAATAAACCTTGCATGACAAACCTGCAAAATATTTTATCAGCTGTTATTGGAAAGTGATTTT
AAGCAATTGCTTCCTCAGTGTCAGGGCACATGTGAATTTCCACACCAAACAGAGCATGAGGAACCAGTTG
ACATGCTGGGTTGTGACTGGCAGCTTTAGCAGCCTCGGTACTGAAGCCACACCAGTGTCCGGATGGAAGT
CTGCATCTGAGGTTGCTCAGTGTCCCGGTCATTCATTTACACATTTTAACTTGCATTAAAGAGCTGTTCT
TTTCTGTGGCCTAGACTCTTTTCACTGATCTCAAAATAAACTGGTTTTTTTCAAAAAAAAAAAAAAAACA
AAAACAAAAAAAAAACACAAAAGCTGCATGTCTAAAATTACATGGAGTTAGTGTCTATTCTTTTTCCCCT
TTTGCAGCAACTTACACAGCATTTTTAACACCTTTTTTTTCTAGTTTTTTTGTTCGGTTTTGTTTTCCAT
CAGGAATTTGAGTTCTCTCTAACCCAGCTTACTGTGGGACATAGGAAAACTCAGTAGAAATACCTTTGGT
GATCTTGTTGAGTTTAAGTCTGATCTTGATCTTAAACTCAGTAAGCCACTATCTGCAATTTTGTACATTA
TATAGTATTTTGAAGATATGGAACCTTATGAAAAAAAAATAGCAAATTAGTTCTTTTTCCCCCAGAGGGG
AAAGTTATGTTCTGCAAATAGTGTGTGTCTTATTTTACTGTTGAACAGCAATTGCTATTTATTTTTTTAT
TGCCTAGAACTTCAACATGTTGTATAGGAATCCTGTAGTGCCACTAGTTAAATGCCGAATTCTCATCTGG
ATGTTACCATCAAACATCAGTACACTTGTCATTTCACATGTGTTTAATGTGACAGTTTTTCAGTACTGTA
TGTGTTAATTTCTACTTTTTTTAATATTTAAAATTGCTTTTAAATAAACATATTCTCAGTTGATCCC

Human DDEF1 protein sequence - var1 (public gi: 31873728) (SEQ ID NO: 233)
ETKFTKIEKEKREHAKQHGMIRTEITGAEIAEEMEKERRLFQLQMCEYLIKVNEIKTKKGVDLLQNLIKY
YHAQCNFFQDGLKTADKLKQYIEKLAADLYNIKQTQDEEKKQLTALRDLIKSSLQLDQKEDSQSRQGGYS
MHQLQGNKEYGSEKKGYLLKKSDGIRKVWQRRKCSVKNGILTISHATSNRQPAKLNLLTCQVKPNAEDKK
SFDLISHNRTYHFQAEDEQDYVAWISVLTNSKEEALTMAFRGEQSAGENSLEDLTKAIIEDVQRLPGNDI
CCDCGSSEPTWLSTNLGILTCIECSGIHREMGVHISRIQSLELDKLGTSELLLAKNVGNNSFNDIMEANL
PSPSPKPTPSSDMTVRKEYITAKYVDHRFSRKTCSTSSAKLNELLEAIKSRDLLALIQVYAEGVELMEPL
LEPGQELGETALHLAVRTADQTSLHLVDFLVQNCGNLDKQTALGNTVLHYCSMYSKPECLKLLLRSKPTV
DIVNQAGETALDIAKRLKATQCEDLLSQAKSGKFNPHVHVEYEWNLRQEEIDESDDDLDDKPSPIKKERS
PRPQSFCHSSSISPQDKLALPGFSTPRDKQRLSYGAFTNQIFVSTSTDSPTSPTTEAPPLPPRNAGKGPT
GPPSTLPLSTQTSSGSSTLSKKRPPPPPPGHKRTLSDPPSPLPHGPPNKGAVPWGNDGGPSSSSKTTNKF
EGLSQQSSTSSAKTALGPRVLPKLPQKVALRKTDHLSLDKATIPPEIFQKSSQLAELPQKPPPGDLPPKP
TELAPKPQIGDLPPKPGELPPKPQLGDLPPKPQLSDLPPKPQMKDLPPKPQLGDLLAKSQTGDVSPKAQQ
PSEVTLKSHPLDLSPNVQSRDAIQKQASEDSNDLTPTLPETPVPLPRKINTGKNKVRRVKTIYDCQADND
DELTFIEGEVIIVTGEEDQEWWIGHIEGQPERKGVFPVSFVHILSD

Human DDEF1 protein sequence - var2 (public gi: 6330854) (SEQ ID NO: 234)
KREHAKQHGMIRTEITGAEIAEEMEKERRLFQLQMCEYLIKVNEIKTKKGVDLLQNLIKYYHAQCNFFQD
GLKTADKLKQYIEKLAADLYNIKQTQDEEKKQLTALRDLIKSSLQLDQKESRRDSQSRQGGYSMHQLQGN
KEYGSEKKGYLLKKSDGIRKVWQRRKCSVKNGILTISHATSNRQPAKLNLLTCQVKPNAEDKKSFDLISH
NRTYHFQAEDEQDYVAWISVLTNSKEEALTMAFRGEQSAGENSLEDLTKAIIEDVQRLPGNDICCDCGSS
EPTWLSTNLGILTCIECSGIHREMGVHISRIQSLELDKLGTSELLLAKNVGNNSFNDIMEANLPSPSPKP
TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSAKLNELLEAIKSRDLLALIQVYAEGVELMEPLLEPGQEL
GETALHLAVRTADQTSLHLVDFLVQNCGNLDKQTALGNTVLHYCSMYSKPECLKLLLRSKPTVDIVNQAG
ETALDIAKRLKATQCEDLLSQAKSGKFNPHVHVEYEWNLRQEEIDESDDDLDDKPSPIKKERSPRPQSFC
HSSSISPQDKLALPGFSTPRDKQRLSYGAFTNQIFVSTSTDSPTSPTTEAPPLPPRNAGKGPTGPPSTLP

LSTQTSSGSSTLSKKRPPPPPPGHKRTLSDPPSPLPHGPPNKGAVPWGNDGGPSSSSKTTNKFEGLSQQS
STSSAKTALGPRVLPKLPQKVALRKTDHLSLDKATIPPEIFQKSSQLAELPQKPPPGDLPPKPTELAPKP
QIGDLPPKPGELPPKPQLGDLPPKPQLSDLPPKPQMKDLPPKPQLGDLLAKSQTGDVSPKAQQPSEVTLK
SHPLDLSPNVQSRDAIQKQASEDSNDLTPTLPETPVPLPRKINTGKNKVRRVKTIYDCQADNDDELTFIE
GEVIIVTGEEDQEWWIGHIEGQPERKGVFPVSFVHILSD

Human DDEF1 protein sequence - var3 (public gi: 7689054) (SEQ ID NO: 235)
MNAHLSDVLKHPLHPFFIIFLVLFLDKFKXXKRLFKMNLKVNPSHSFPTVCRIQVLKPSVSFPMLFVKLQ
L

Human DDEF1 protein sequence - var4 (public gi: 18088818) (SEQ ID NO: 236)
MNAHLSDVLKHPLHPFFIIFLVLFLDKFKLLKDYSR

Human DDEF1 protein sequence - var5 (Predicted by Proteologics) (SEQ ID NO: 237)
MIGQPQEACRSHHKSHKALDQDRTALQKVKKSVKAIYNSGQDHVQNEENYAQVLDKFGSNFLSRDNPDLG
TAFVKFSTLTKELSTLLKNLLQGLSHNVIFTLDSLLKGDLKGVKGDLKKPFDKAWKDYETKFTKIEKEKR
EHAKQHGMIRTEITGAEIAEEMEKERRLFQLQMCEYLIKVNEIKTKKGVDLLQNLIKYYHAQCNFFQDGL
KTADKLKQYIEKLAADLYNIKQTQDEEKKQLTALRDLIKSSLQLDQKESRRDSQSRQGGYSMHQLQGNKE
YGSEKKGYLLKKSDGIRKVWQRRKCSVKNGILTISHATSNRQPAKLNLLTCQVKPNAEDKKSFDLISHNR
TYHFQAEDEQDYVAWISVLTNSKEEALTMAFRGEQSAGENSLEDLTKAIIEDVQRLPGNDICCDCGSSEP
TWLSTNLGILTCIECSGIHREMGVHISRIQSLELDKLGTSELLLAKNVGNNSFNDIMEANLPSPSPKPTP
SSDMTVRKEYITAKYVDHRFSRKTCSTSSAKLNELLEAIKSRDLLALIQVYAEGVELMEPLLEPGQELGE
TALHLAVRTADQTSLHLVDFLVQNCGNLDKQTALGNTVLHYCSMYSKPECLKLLLRSKPTVDIVNQAGET
ALDIAKRLKATQCEDLLSQAKSGKFNPHVHVEYEWNLRQEEIDESDDDLDDKPSPIKKERSPRPQSFCHS
SSISPQDKLALPGFSTPRDKQRLSYGAFTNQIFVSTSTDSPTSPTTEAPPLPPRNAGKGPTGPPSTLPLS
TQTSSGSSTLSKKRPPPPPPGHKRTLSDPPSPLPHGPPNKGAVPWGNDGGPSSSSKTTNKFEGLSQQSST
SSAKTALGPRVLPKLPQKVALRKTDHLSLDKATIPPEIFQKSSQLAELPQKPPPGDLPPKPTELAPKPQI
GDLPPKPGELPPKPQLGDLPPKPQLSDLPPKPQMKDLPPKPQLGDLLAKSQTGDVSPKAQQPSEVTLKSH
PLDLSPNVQSRDAIQKQASEDSNDLTPTLPETPVPLPRKINTGKNKVRRVKTIYDCQADNDDELTFIEGE
VIIVTGEEDQEWWIGHIEGQPERKGVFPVSFVHILSD

Human DDEF1 pray sequence - var1 (SEQ ID NO: 54)
GCGCCGCCATGGTAGTACCCATACGACGTACCAGTATTACGCTCATATGGCCATGGCAGGCCAGTGAATT
CCACACCAAGCAGTGGTATCAACGCAGAGTGGGCACAAAAGCCACGCACGCTGGANGACCTGCCCCCAAC
AGCCCACAGAAACTGGCCCCCAAGCCCCAAATTGGAGATTTGCCGCCTAAGCCAGGAGAACTGCCCCCCA
AACCACAGCTGGGGGACCTGCCACCCAAACCCCAACACTCAGACTTACCTCCCAAACCACAGATGAAGGA
CCTGCCCCCCAAACCACAGCTGGGAGACCTGCTAGCAAAATCCCAGACTGGAGATGTCTCACCCAAGGCT
CAGCAACCCTCTGAGGTCACACTGAAGTCACACCCATTGGATCTATCCCCAAATGTGCAGTCCAGAGACG
CCATCCAAAAGCAAGCATNTGAAGACTCCAACGACCTCACGCCTACTCTGCCAGAGACGCCCGTACCACT
GCCCANAAAAATCANTACGGGGAAAANTAANNTGAGGCGAGTGAAAACCTTTAATGACTGCCAGGCANAC
ANNATGACAAGCTCNATTCNTCNAGGGANAAGTGTTATCGTNCAGGGAAAAAGNNCNGGATTGTGGGTCC
NNCAATTTTCNTCCNNTNNTCNNACTTATTANAATNGCNNGGCAGGNNCCAATNGAACNCCNAANNNGNN
GAAAANAGGNNTTTNNNCAAGGANCNTNNNNNTNGTTTNTTCCCNAAANNTTNNTTNGGNNTTTTTTTNC
NCNCNTTTTTNTNNAAAAACNCNGNANNNNNNNNCAAGGNNNCCCNTNTNTNNCNTTNGGGGGGGGGNNG
NNTNNNGGGGGGNNNANACCCCCCC

<u>Unigene Name</u>: EIF3S3 <u>Unigene ID</u>: Hs.58189 <u>Clone ID</u>: 3GD_18

Human EIF3S3 mRNA sequence - var1 (public gi: 2351379) (SEQ ID NO: 55)
GAAAGATGGCGTCCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCGCCGGCGCAGC
AGGGAAAGGCAAAGGCAAAGGCGGCTCGGGAGATTCAGCCGTGAAGCAAGTGCAGATAGATGGCCTTGTG
GTATTAAAGATAATCAAACATTATCAAGAAGAAGGACAAGGAACTGAAGTTGTTCAAGGAGTGCTTTTGG
GTCTGGTTGTAGAAGATCGGCTTGAAATTACCAACTGCTTTCCTTTCCCTCAGCACACAGAGGATGATGC
TGACTTTGATGAAGTCCAATATCAGATGGAAATGATGCGGAGCCTTCGCCATGTAAACATTGATCATCTT
CACGTGGGCTGGTATCAGTCCACATACTATGGCTCATTCGTTACCCGGGCACTCCTGGACTCTCAGTTTA
GTTACCAGCATGCCATTGAAGAATCTGTCGTTCTCATTTATGATCCCATAAAAACTGCCCAAGGATCTCT
CTCACTAAAGGCATACAGACTGACTCCTAAACTGATGGAAGTTTGTAAAGAAAAGGATTTTTCCCCTGAA
GCATTGAAAAAAGCAAATATCACCTTTGAGTACATGTTTGAAGAAGTGCCGATTGTAATTAAAAATTCAC
ATCTGATCAATGTCCTAATGTGGGAACTTGAAAAGAAGTCAGCTGTTGCAGATAAACATGAATTGCTCAG

CCTTGCCAGCAGCAATCATTTGGGGAAGAATCTACAGTTGCTGATGGACAGAGTGGATGAAATGAGCCAA
GATATAGTTAAATACAACACATACATGAGGAATACTAGTAAACAACAGCAGCAGAAACATCAGTATCAGC
AGCGTCGCCAGCAGGAGAATATGCAGCGCCAGAGCCGAGGAGAACCCCCGCTCCCTGAGGAGGACCTGTC
CAAACTCTTCAAACCACCACAGCCGCCTGCCAGGATGGACTCGCTGCTCATTGCAGGCCAGATAAACACT
TACTGCCAGAACATCAAGGAGTTCACTGCCCAAAACTTAGGCAAGCTCTTCATGGCCCAGGCTCTTCAAG
AATACAACAACTAAGAAAAGGAAGTTTCCAGAAAAGAAGTTAACATGAACTCTTGAAGTCACACCAGGGC
AACTCTTGGAAGAAATATATTTGCATATTGAAAAGCACAGAGGATTTCTTTAGTGTCATTGCCGATTTTG
GCTATAACAGTGTCTTTCTAGCCATAATAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAA

Human EIF3S3 mRNA sequence - var2 (public gi: 21751901) (SEQ ID NO: 56)
AGGCGCGTAGCAAGAGCTTCTCTGAAAGACTGGCAGTATAGTAGTAGTCAGTGATAATATTGAGCCTTAA
TATGTTCCAGACACTGTCCTAAGTGATTTACCTTACATTATTCCCCTGAATGTTTATAATTCCCAAGTGA
AAGAAGGAAATGATATATTGGATAGCTATGAGTGGGGAGGTTTGTACTGGCTGCTTTCCCATAAAGAAAT
TAAGCACGTTCACGAAGGGCACGTAGTTTGTTAGTGTCTGGAACCCAGTTTTCGTGCCTGAAGTTCAAAT
GTTCTTGCTACACCACCATAGAAACTAACGTCACTCAGGAACCATTTGTCAGGGCAAAGGGTGCCACCAT
TTTGCATTTCCTCCTGCCTTAGGACCATCCTAAATCACTCGCATGGAGTGTTTTGGAAGAACTCTCAAGA
GCTTCGTTTGCCTAGAGTCAGAATTCCTAACCTTGAGTCCTGGTTTTGCCACAAACCCAAGCCGTTTGAT
CTTGGGCAACTCCCGAAGAAAGCTGGGTTCAACTTCCTCACTGTCAAACTGGTTGTAGGTCTAGATAAGT
TTCAAGTACTCTTTTTATGTGCATGGTCTCTGACATAGGAAGACTACATACTGGGCCAGTAACAGGAAGG
CACAAAGCTGACTGGAGGTTTAAAAATTACTTGGTCAATTTGATTAATGAGGAGAATGAATCAGAAAATT
TCAAGTTCTCCCGTGGCTAACTGTCGAGTATCCACTTCAAGATCATTCCATCGGAAAGAGGTCGAAAATG
TACAGTAGGCATGCACAAAGGATACCGCCTGGAAAGAAGATGGCGTCCCGCAAGGAAGGTACCGGCTCTA
CTGCCACCTCTTCCAGCTCCACCGCCGGCGCAGCAGGGAAAGGCAAAGGCAAAGGCGGCTCGGGAGATTC
AGCCGTGAAGCAAGTGCAGATAGATGGCCTTGTGGTATTAAAGATAATCAAACATTATCAAGAAGAAGGA
CAAGGAACTGAAGTTGTTCAAGGAGTGCTTTTGGGTCTGGTTGTAGAAGATCGGCTTGAAATTACCAACT
GCTTTCCTTTCCCTCAGCACACAGAGGATGATGCTGACTTTGATGAAGTCCAATATCAGATGGAAATGAT
GCGGAGCCTTCGCCATGTAAACATTGATCATCTTCACGTGGGCTGGTATCAGTCCACATACTATGGCTCA
TTCGTTACCCGGGCACTCCTGGACTCTCAGTTTAGTTACCAGCATGCCATTGAAGAATCTGTCGTTCTCA
TTTATGATCCCATAAAAACTGCCCAAGGATCTCTCTCACTAAAGGCATACAGACTGACTCCTAAACTGAT
GGAAGTTTGTAAAGAAAAGGATTTTTCCCCTGAAGCACTGAAAAAAGCAAATATCACCTTTGAGTACATG
TTTGAAGAAGTGCCGATTGTAATTAAAAATTCACATCTGATCAATGTCCTAATGTGGGAACTTGAAAAGA
AGTCAGCTGTTGCAGATAAACATGAATTGCTCAGCCTTGCCAGCAGCAATCATTTGGGGAAGAATCTACA
GTTGCTGATGGACAGAGTGGATGAAATGAGCCAAGATATAGTTAAATACAACACATACATGAGGAATACT
AGTAAACAACAGCAGCAGAAACATCAGTATCAGCAGCGTCGCCAGCAGGAGAATATGCAGCGCCAGAGCC
GAGGAGAACCCCCGCTCCCTGAGGAGGACCTGTCCAAACTCTTCAAACCACCACAGCCGCCTGCCAGGAT
GGACTCGCTGCTCATTGCAGGCCAGATAAACACTTACTGCCAGAACATCAAGGAGTTCACTGCCCAAAAC
TTAGGCAAGCTCTTCATGGCCCAGGCTCTTCAAGAATACAACAACTAAGAAAAGGAAGTTTCCAGAAAAG
AAGTTAACATGAACTCTTGAAGTCACACCAGGGCAACTCTTGGAAGAAATATATTTGCATATTGAAAAGC
ACAGAGGATTTCTTTAGTGTCATTGCCGATTTTGGCTATAACAGTGTCTTTCTAGCCATAATAAAATAAA
ACAAAATCTTG

Human EIF3S3 mRNA sequence - var3 (public gi: 12653234) (SEQ ID NO: 57)
GGCACGAGGATGGCGTCCCGCAAGGAAGGTACCGGCTCTACTGCCACCTCTTCCAGCTCCACCGCCGGCG
CAGCAGGGAAAGGCAAAGGCAAAGGCGGCTCGGGAGATTCAGCCGTGAAGCAAGTGCAGATAGATGGCCT
TGTGGTATTAAAGATAATCAAACATTATCAAGAAGAAGGACAAGGAACTGAAGTTGTTCAAGGAGTGCTT
TTGGGTCTGGTTGTAGAAGATCGGCTTGAAATTACCAACTGCTTTCCTTTCCCTCAGCACACAGAGGATG
ATGCTGACTTTGATGAAGTCCAATATCAGATGGAAATGATGCGGAGCCTTCGCCATGTAAACATTGATCA
TCTTCACGTGGGCTGGTATCAGTCCACATACTATGGCTCATTCGTTACCCGGGCACTCCTGGACTCTCAG
TTTAGTTACCAGCATGCCATTGAAGAATCTGTCGTTCTCATTTATGATCCCATAAAAACTGCCCAAGGAT
CTCTCTCACTAAAGGCATACAGACTGACTCCTAAACTGATGGAAGTTTGTAAAGAAAAGGATTTTTCCCC
TGAAGCATTGAAAAAAGCAAATATCACCTTTGAGTACATGTTTGAAGAAGTGCCGATTGTAATTAAAAAT
TCACATCTGATCAATGTCCTAATGTGGGAACTTGAAAAGAAGTCAGCTGTTGCAGATAAACATGAATTGC
TCAGCCTTGCCAGCAGCAATCATTTGGGGAAGAATCTACAGTTGCTGATGGACAGAGTGGATGAAATGAG
CCAAGATATAGTTAAATACAACACATACATGAGGAATACTAGTAAACAACAGCAGCAGAAACATCAGTAT
CAGCAGCGTCGCCAGCAGGAGAATATGCAGCGCCAGAGCCGAGGAGAACCCCCGCTCCCTGAGGAGGACC

TGTCCAAACTCTTCAAACCACCACAGCCGCCTGCCAGGATGGACTCGCTGCTCATTGCAGGCCAGATAAA
CACTTACTGCCAGAACATCAAGGAGTTCACTGCCCAAAACTTAGGCAAGCTCTTCATGGCCCAGGCTCTT
CAAGAATACAACAACTAAGAAAAGGAAGTTTCCAGAAAAGAAGTTAACATGAACTCTTGAAGTCACACCA
GGGCAACTCTTGGAAGAAATATATTTGCATATTGAAAAGCACAGAGGATTTCTTTAGTGTCATTGCCGAT
TTTGGCTATAACAGTGTCTTTCTAGCCATAATAAAATAAAACAAAATCTTGAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human EIF3S3 protein sequence - var1 (public gi: 12653235) (SEQ ID NO: 238)
MASRKEGTGSTATSSSSTAGAAGKGKGKGGSGDSAVKQVQIDGLVVLKIIKHYQEEGQGTEVVQGVLLGL
VVEDRLEITNCFPFPQHTEDDADFDEVQYQMEMMRSLRHVNIDHLHVGWYQSTYYGSFVTRALLDSQFSY
QHAIEESVVLIYDPIKTAQGSLSLKAYRLTPKLMEVCKEKDFSPEALKKANITFEYMFEEVPIVIKNSHL
INVLMWELEKKSAVADKHELLSLASSNHLGKNLQLLMDRVDEMSQDIVKYNTYMRNTSKQQQQKHQYQQR
RQQENMQRQSRGEPPLPEEDLSKLFKPPQPPARMDSLLIAGQINTYCQNIKEFTAQNLGKLFMAQALQEY
NN

Unigene Name: EPS8L2 **Unigene ID:** Hs.55016

Human EPS8L2 mRNA sequence - var1 (public gi: 21264615) (SEQ ID NO: 58)
GTCGACGGCCATTACCAATCGCGAAACCCCGCAACCTGTCGCTCAGGTTCCTCCTCTCCCGGCCCCGCCC
CGGCCCGGCCCCGCCGAGCGTCCCACCCGCCCGCGGGAGACCTGGCGCCCCGGCCGAGGCGCGAACAGAC
GGACGCACCGGCGAGCGCCGAGGGGACAGGCCGAGCGCGGGGCGCCGGAGGCAGGTGTGGGACAGGCACT
GGCCTCAGACCGGGGCCACACTGAGGTCTGCCCTTCTCCCGCTGGCCGCCACCCAAGACACCATGAGCCA
GTCCGGGGCCGTGAGCTGCTGCCCGGGTGCCACCAATGGCAGCCTGGGCCGGTCCGACGGTGTGGCCAAG
ATGAGCCCCAAGGACCTGTTTGAGCAGAGGAAGAAGTATTCCAACTCCAACGTCATCATGCACGAGACCT
CGCAGTACCACGTCCAGCACCTGGCCACATTCATCATGGACAAGAGCGAAGCCATCACGTCTGTGGACGA
CGCCATCCGGAAGCTGGTGCAGCTGAGCTCCAAGGAGAAGATCTGGACCCAGGAGATGCTGCTGCAGGTG
AACGACCAGTCGTTGCGGCTGCTGGACATCGAGTCACAGGAGGAGCTGGAAGACTTCCCGCTGCCCACGG
TGCAGCGCAGCCAGACGGTCCTCAACCAGCTGCGCTACCCGTCTGTGCTGCTGCTCGTGTGCCAGGACTC
GGAGCAGAGCAAGCCGGATGTCCACTTCTTCCACTGCGATGAGGTGGAGGCAGAGCTGGTGCACGAGGAC
ATCGAGAGCGCGTTGGCCGACTGCCGGCTGGGCAAGAAGATGCGGCCGCAGACCCTGAAGGGACACCAGG
AGAAGATTCGGCAGCGGCAGTCCATCCTGCCTCCTCCCCAGGGCCCGGCGCCCATCCCCTTCCAGCACCG
CGGCGGGGATTCCCCGGAGGCCAAGAATCGCGTGGGCCCGCAGGTGCCACTCAGCGAGCCAGGTTTCCGC
CGTCGGGAGTCGCAGGAGGAGCCGCGGGCCGTGCTGGCTCAGAAGATAGAGAAGGAGACGCAAATCCTCA
ACTGCGCCCTGGACGACATCGAGTGGTTTGTGGCCCGGCTGCAGAAGGCAGCCGAGGCTTTCAAGCAGCT
GAACCAGCGGAAAAAGGGGAAGAAGAAGGGCAAGAAGGCGCCAGCAGAGGGCGTCCTCACACTGCGGGCA
CGGCCCCCCTCTGAGGGCGAGTTCATCGACTGCTTCCAGAAAATCAAGCTGGCGATTAACTTGCTGGCAA
AGCTGCAGAAGCACATCCAGAACCCCAGCGCCGCGGAGCTCGTGCACTTCCTCTTCGGGCCTCTGGACCT
GATCGTCAACACCTGCAGTGGCCCAGACATCGCACGCTCCGTCTCCTGCCCACTGCTCTCCCGAGATGCC
GTGGACTTCCTGCGCGGCCACCTGGTCCCTAAGGAGATGTCGCTGTGGGAGTCACTGGGAGAGAGCTGGA
TGCGGCCCCGTTCCGAGTGGCCGCGGGAGCCACAGGTGCCCCTCTACGTGCCCAAGTTCCACAGCGGCTG
GGAGCCTCCTGTGGATGTGCTGCAGGAGGCCCCCTGGGAGGTGGAGGGGCTGGCGTCTGCCCCCATCGAG
GAGGTGAGTCCAGTGAGCCGACAGTCCATAAGAAACTCCCAGAAGCACAGCCCCACTTCAGAGCCCACCC
CCCCGGGGGATGCCCTACCACCAGTCAGCTCCCCACATACTCACAGGGGCTACCAGCCAACACCAGCCAT
GGCCAAGTACGTCAAGATCCTGTATGACTTCACAGCCCGAAATGCCAACGAGCTATCGGTGCTCAAGGAT
GAGGTCCTAGAGGTGCTGGAGGACGGCCGGCAGTGGTGGAAGCTGCGCAGCCGCAGCGGCCAGGCGGGGT
ACGTGCCCTGCAACATCCTAGGCGAGGCGCGACCGGAGGACGCCGGCGCCCCGTTCGAGCAGGCCGGTCA
GAAGTACTGGGGCCCCGCCAGCCCGACCCACAAGCTACCCCCAAGCTTCCCGGGGAACAAAGACGAGCTC
ATGCAGCACATGGACGAGGTCAACGACGAGCTCATCGGAAAAATCAGCAACATCAGGGCGCAGCCACAGA
GGCACTTCCGCGTGGAGCGCAGCCAGCCCGTGAGCCAGCCGCTCACCTACGAGTCGGGTCCGGACGAGGT
CCGCGCCTGGCTGGAAGCCAAGGCCTTCAGCCCGCGGATCGTGGAGAACCTGGGCATCCTGACCGGGCCG
CAGCTCTTCTCCCTCAACAAGGAGGAGCTGAAGAAAGTGTGCGGCGAGGAGGGCGTCCGCGTGTACAGCC
AGCTCACCATGCAGAAGGCCTTCCTGGAGAAGCAGCAAAGTGGGTCGGAGCTGGAAGAACTCATGAACAA
GTTTCATTCCATGAATCAGAGGAGGGGGGAGGACAGCTAGGCCCAGCTGCCTTGGGCTGGGGCCTGCGGA
GGGGAAGCCCACCCACAATGCATGGAGTATTATTTTTATATGTGTATGTATTTTGTATCAAGGACACGGA
GGGGGTGTGGTGCTGGCTAGAGGTCCCTGCCCCTGTCTGGAGGCACAACGCCCATCCTTAGGCCAAACAG
TACCCAAGGCCTCAGCCCACACCAAGACTAATCTCAGCCAAACCTGCTGCTTGGTGGTGCCAGCCCCTTG
TCCACCTTCTCTTGAGGCCACAGAACTCCCTGGGGCTGGGGCCTCTTTCTCTGGCCTCCCCTGTGCACCT
GGGGGGTCCTGGCCCCTGTGATGCTCCCCCATCCCCACCCACTTCTACATCCATCCACACCCCAGGGTGA
GCTGGAGCTCCAGGCTGGCCAGGCTGAACCTCGCACACACGCAGAGTTCTGCTCCCTGAGGGGGGCCCGG
GAGGGGCTCCAGCAGGAGGCCGTGGGTGCCATTCGGGGGAAAGTGGGGAACGACACACACTTCACCTGC

AAGGGCCGACAACGCAGGGGACACCGTGCCGGCTTCAGACACTCCCAGCGCCCACTCTTACAGGCCCAGG
ACTGGAGCTTTCTCTGGCCAAGTTTCAGGCCAATGATCCCCGCATGGTGTTGGGGTGCTGGTGTGTCTT
GGTGCCTGGACTTGAGTCTCACCCTACAGATGAGAGGTGGCTGAGGCACCAGGGCTAAGCAATTAAACCA
GTTAAGTCTCCCAGGAAAAAAAAAAAAAAAA

Human EPS8L2 protein sequence - var1 (public gi: 21264616) (SEQ ID NO: 239)
MSQSGAVSCCPGATNGSLGRSDGVAKMSPKDLFEQRKKYSNSNVIMHETSQYHVQHLATFIMDKSEAITS
VDDAIRKLVQLSSKEKIWTQEMLLQVNDQSLRLLDIESQEELEDFPLPTVQRSQTVLNQLRYPSVLLLVC
QDSEQSKPDVHFFHCDEVEAELVHEDIESALADCRLGKKMRPQTLKGHQEKIRQRQSILPPPQGPAPIPF
QHRGGDSPEAKNRVGPQVPLSEPGFRRRESQEEPRAVLAQKIEKETQILNCALDDIEWFVARLQKAAEAF
KQLNQRKKGKKKGKKAPAEGVLTLRARPPSEGEFIDCFQKIKLAINLLAKLQKHIQNPSAAELVHFLFGP
LDLIVNTCSGPDIARSVSCPLLSRDAVDFLRGHLVPKEMSLWESLGESWMRPRSEWPREPQVPLYVPKFH
SGWEPPVDVLQEAPWEVEGLASAPIEEVSPVSRQSIRNSQKHSPTSEPTPPGDALPPVSSPHTHRGYQPT
PAMAKYVKILYDFTARNANELSVLKDEVLEVLEDGRQWWKLRSRSGQAGYVPCNILGEARPEDAGAPFEQ
AGQKYWGPASPTHKLPPSFPGNKDELMQHMDEVNDELIRKISNIRAQPQRHFRVERSQPVSQPLTYESGP
DEVRAWLEAKAFSPRIVENLGILTGPQLFSLNKEELKKVCGEEGVRVYSQLTMQKAFLEKQQSGSELEEL
MNKFHSMNQRRGEDS

Human EPS8L2 pray sequence - var1 (SEQ ID NO: 59)
TCNTNCGCCGCCATGGNAGTACCCATACGACGTACCAGNATTACGCTCATATGGCCATGGNAGGCCAGTG
AATTCCACCCAAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCGGGGGAACAAAGACGAGCTCATGC
AGCACATGGACGAGGTCAACGACGAGCTCATCCGGAAAATCAGCAACATCAGGGCGCAGCCACAGAGGCA
CTTCCGCGTGGAGCGCAGCCAGCCCGTGAGCCAGCCGCTCACCTACGAGTCGGGTCCGGACGAGGTCCGC
GCCTGGCTGGAAGCCAAGGCCTTCAGCCCGCGGATCGTGGAGAACCTGGGCATCCTGACCGGGCCGCAGC
TCTTCTCCCTCAACAAGGAGGAGCTGAAGAAAGTGTGCGGCGAGGAGGGCGTCCGCGTGTACAGCCAGCT
CACCATGCAGAAGGCCTTCCTGGAGAAGCAGCAAAGTGGGTCGGAGCTGGAAGAACTCATGAACAAGTTT
CATTCCATGAATCAGAGGAGGGGGGAGGACAGCTAGGCCCAGCTGCCTTGGGCTGGGGCCTGCGGAGGGG
AAGCCCACCCACAATGCATGGAGTATTATTTTTATATGTGTATGTATTTTGTATCAAGGACACGGAGGGG
GGTGTGGTGCTTGGCTANAGGTCCCTGCCCCTGTTTGGNAGGCACAACNCCCATNCTTTAGNCCAAANAG
TNACCCAANGGCCTNAACCCCAANCAAGNTTATTTTNANNCCAAACNNGNTTGNTTGGTTGGTNCCAACC
CCNTTGTTGGTGCCNNNCCCNTTGTNCANCNTTNNTTTTNGGNCNCNANAANTNCCTNGGGGTNGGGGGN
CNTTTTTTNTNN

Human EPS8L2 pray sequence - var2 (SEQ ID NO: 60)
CGAGCGCCGCCTGGNATACCCATACGACGTACCAGNATTACGCTCATATGGCCATGGNAGGCCAGTGAAT
TCCACCCAAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCGGGGGAACAAAGACGAGCTCATGCAGC
ACATGGACGAGGTCAACGACGAGCTCATCCGGAAAATCAGCAACATCAGGGCGCAGCCACAGAGGCACTT
CCGCGTGGAGCGCAGCCAGCCCGTGAGCCAGCCGCTCACCTACGAGTCGGGTCCGGACGAGGTCCGCGCC
TGGCTGGAAGCCAAGGCCTTCAGCCCGCGGATCGTGGAGAACCTGGGCATCCTGACCGGGCCGCAGCTCT
TCTCCCTCAACAAGGAGGAGCTGAAGAAAGTGTGCGGCGAGGAGGGCGTCCGCGTGTACAGCCAGCTCAC
CATGCAGAAGGCCTTCCTGGAGAAGCAGCAAAGTGGGTCGGAGCTGGAAGAACTCATGAACAAGTTTCAT
TCCATGAATCAGAGGAGGGGGGAGGACAGCTAGGCCCAGCTGCCTTGGGCTGGGGCCTGCGGAGGGGAAG
CCCACCCACAATGCATGGAGTATTATTTTTATATGTGTATGTATTTTGTATCAAGGACACGGAGGGGGTG
TGGTGCTGGCTANAGTTCCCTGCCCTTGTNTGGAGGCACACNCCCATCCTTAGGCCAAACANTACCNAGG
NCTNANCCCACACCAANACTATTTTAACCNAACTNGNTGNTTGGTGGTGCCNNCCCNTTGGTGNTNCCNC
CCNTTNTCCCNTTTTTTTGNGNCCNAAAATTCNTGGGGCTGGGCNTTTTTTTTTGGCNNCCCTTNNNNCN
TNGGGGTTCTGGNCCNTNNNNTNTNCCCCTNCCCCCCNTTTTNNNTNTTN

Human GOCAP1 mRNA sequence - var1 (public gi: 10438060) (SEQ ID NO: 61)
GATACGTGGCTGCCGTCTGTCCCCGCTGAGGAGGTGCAGCAGCCGGAGATGGCGGCGGTGCTGAACGCAG
AGCGACTCGAGGTGTCCGTCGACGGCCTCACGCTCAGCCCGGACCCGGAGGAGCGGCCTGGGGCGGAGGG
CGCCCCGCTGCTGCCGCCACCGCTGCCACCGCCCTCGCCACCTGGATCCGGTCGCGGCCCGGGCGCCTCA
GGGGAGCAGCCCGAGCCCGGGGAGGCGGCGGCTGGGGGCGCGGCGGAGGAGGCGCGGCGGCTGGAGCAGC
GCTGGGGTTTCGGCCTGGAGGAGTTGTACGGCCTGGCACTGCGCTTCTTCAAAGAAAAAGATGGCAAAGC
ATTTCATCCAACTTATGAAGAAAAATTGAAGCTTGTGGCACTGCATAAGCAAGTTCTTATGGGCCCATAT
AATCCAGACACTTGTCCTGAGGTTGGATTCTTTGATGTGTTGGGGAATGACAGGAGGAGAGAATGGGCAG
CCCTGGGAAACATGTCTAAAGAGGATGCCATGGTGGAGTTTGTCAAGCTCTTAAATAGGTGTTGCCATCT
CTTTTCAACATATGTTGCGTCCCACAAAATAGAGAAGGAAGAGCAAGACAAAAAAAGGAAGGAGGAAGAG
GAGCGAAGGCGGCGTGAAGAGGAAGAAAGAGAACGTCTGCAAAAGGAGGAAGAGAAACGTAGGAGAGAAG
AAGAGGAAAGGCTTCGACGGGAGGAAGAGGAAAGGAGACGGATAGAAGAAGAAAGGCTTCGGTTGGAGCA

GCAAAAGCAGCAGATAATGGCAGCTTTAAACTCCCAGACTGCCGTGCAGTTCCAGCAGTATGCAGCCCAA
CAGTATCCAGGGAACTACGAACAGCAGCAAATTCTCATCCGCCAGTTGCAGGAGCAACACTATCAGCAGT
ACATGCAGCAGTTGTATCAAGTCCAGCTTGCACAGCAACAGGCAGCATTACAGAAACAACAGGAAGTAGT
AGTGGCTGGGTCTTCCTTGCCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAGTT
AATGGACAGGCCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAACCAGAAGCTGCAGAAGAAGCCC
TGGAGAATGGACCAAAAGAATCTCTTCCAGTAATAGCAGCTCCATCCATGTGGACACGACCTCAGATCAA
AGACTTCAAAGAGAAGATTCAGCAGGATGCAGATTCCGTGATTACAGTGGGCCGAGGAGAAGTGGTCACT
GTTCGAGTACCCACCCATGAAGAAGGATCATATCTCTTTTGGGAATTTGCCACAGACAATTATGACATTG
GGTTTGGGGTGTATTTTGAATGGACAGACTCTCCAAACACTGCTGTCAGCGTGCATGTCAGTGAGTCCAG
CGATGACGACGAGGAGGAAGAAGAAAACATCGGTTGTGAAGAGAAAGCCAAAAAGAATGCCAACAAGCCT
TTGCTGGATGAGATTGTGCCTGTGTACCGACGGGACTGTCATGAGGAGGTGTATGCTGGCAGCCATCAAT
ATCCAGGGAGAGGAGTCTATCTCCTCAAGTTTGACAACTCCTACTCTTTGTGGCGGTCAAAATCAGTCTA
CTACAGAGTCTATTATACTAGATAAAAATGTTGTTACAAAGTCTGGAGTCTAGGGTTGGGCAGAAGATGA
CATTTAATTTGGAAATTTCTTTTTACTTTTGTGGAGCATTAGAGTCACAGTTTACCTTATTGATATTGGT
CTGATGGTTTGTGAACTCTTGCTGGGAATCAAAATTTCCTTGAGACTCTTTAGCATTCATACTTTGGGGT
TAAAGGAGATTCCTCAGACTCATCCAGCCCTTGGGTGCTGACCAGCAGAGTCACTAGTGGATGCTGAAGT
TACATGAGCTACATGTTAAATATTTAAAGTCTCCAAAATAAAACACCCCAACGTTGACCTTACCCGGCTG
ATGGTTAGCCCCTTGCTGCCTGCTCCATGTGTCTTATGAGAGCCCGTAGTTACAGTGTCCTCTAATTTGA
AATCCATAAGTTAACAAGTCTATATCAGGTGCAGCTGGCTTTGATTAAAGGCCATTTTTAAAACTTAAAA
ACTCAACACCTCACAGATTATAAAAAAAAAAAAAAAAAA

Human GOCAP1 mRNA sequence - var2 (public gi: 15826851) (SEQ ID NO: 62)
GGAAGTCGATACGTGGCTGCCTTCTGTCCCCGCTGAGGAGGTGCAGCAGCCGGAGATGGCGGCGGTGCTG
AACGCAGAGCGACTCGAGGTGTCCGTCGACGGCCTCACGCTCAGCCCGGACCCGGAGGAGCGGCCTGGGG
CGGAGGGCGCCCCGCTGCTGCCGCCACCGCTGCCACCGCCCTCGCCACCTGGATCCGGTCGCGGCCCGGG
CGCCTCAGGGGAGCAGCCCGAGCCCGGGGAGGCGGCGGCTGGGGGCGCGGCGGAGGAGGCGCGGCGGCTG
GAGCAGCGCTGGGGTTTCGGCCTGGAGGAGTTGTACGGCCTGGCACTGCGCTTCTTCAAAGAAAAAGATG
GCAAAGCATTTCATCCAACTTATGAAGAAAAATTGAAGCTTGTGGCACTGCATAAGCAAGTTCTTATGGG
CCCATATAATCCAGACACTTGTCCTGAGGTTGGATTCTTTGATGTGTTGGGGAATGACAGGAGGAGAGAA
TGGGCAGCCCTGGGAAACATGTCTAAAGAGGATGCCATGGTGGAGTTTGTCAAGCTCTTAAATAGGTGTT
GCCATCTCTTTTCAACATATGTTGCGTCCCACAAAATAGAGAAGGAAGAGCAAGAAAAAAAAAGGAAGGA
GGAAGAGGAGCGAAGGCGGCGTGAAGAGGAAGAAAGAGAACGTCTGCAAAAGGAGGAAGAGAAACGTAGG
AGAGAAGAAGAGGAAAGGCTTCGACGGGAGGAAGAGGAAAGGAGACGGATAGAAGAAGAAAGGCTTCGGT
TGGAGCAGCAAAAGCAGCAGATAATGGCAGCTTTAAACTCCCAGACTGCCGTGCAGTTCCAGCAGTATGC
AGCCCAACAGTATCCAGGGAACTACGAACAGCAGCAAATTCTCATCCGCCAGTTGCAGGAGCAACACTAT
CAGCAGTACATGCAGCAGTTGTATCAAGTCCAGCTTGCACAGCAACAGGCAGCATTACAGAAACAACAGG
AAGTAGTAGTGGCTGGGTCTTCCTTGCCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGAT
GTCAGTTAATGGACAGGCCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAACCAGAAGCTGCAGAA
GAAGCCCTGGAGAATGGACCAAAAGAATCTCTTCCAGTAATAGCAGCTCCATCCATGTGGACACGACCTC
AGATCAAAGACTTCAAAGAGAAGATTCAGCAGGATGCAGATTCCGTGATTACAGTGGGCCGAGGAGAAGT
GGTCACTGTTCGAGTACCCACCCATGAAGAAGGATCATATCTCTTTTGGGAATTTGCCACAGACAATTAT
GACATTGGGTTTGGGGTGTATTTTGAATGGACAGACTCTCCAAACACTGCTGTCAGCGTGCATGTCAGTG
AGTCCAGCGATGACGACGAGGAGGAAGAAGAAAACATCGGTTGTGAAGAGAAAGCCAAAAAGAATGCCAA
CAAGCCTTTGCTGGATGAGATTGTGCCTGTGTACCGACGGGACTGTCATGAGGAGGTGTATGCTGGCAGC
CATCAATATCCAGGGAGAGGAGTCTATCTCCTCAAGTTTGACAACTCCTACTCTTTGTGGCGGTCAAAAT
CAGTCTACTACAGAGTCTATTATACTAGATAAAAATGTTGTTACAAAGTCTGGAGTCTAGGGTTGGGCAG
AAGATGACATTTAATTTGGAAATTTCTTTTTACTTTTGTGGAGCATTAGAGTCACAGTTTACCTTATTGA
TATTGGTCTGATGGTTTGTGAACTCTTGCTGGGAATCAAAATTTCCTTGAGACTCTTTAGCATTCATACT
TTGGGGTTAAAGGAGATTCCTCAGACTCATCCAGCCCTTGGGTGCTGACCAGCAGAGTCACTAGTGGATG
CTGAAGTTACATGAGCTACATGTTAAATATTTAAAGTCTCCAAAATAAAACACCCCAACGTTGACCTTAC
CCGGCTGATGGTTAGCCCCTTGCTGCCTGCTCCATGTGTCTTATGAGAGCCCGTAGTTACAGTGTCCTCT
AATTTGAAATCCATAAGTTAACAAGTCTATATCAGGTGCATCTGGCTTTGATTAAAGGCCATTTTTAAAA
CTTAAAAACTCAACACCTCACAGATTATAATAGAAAAAGAAATGGCCTCAGTTTGATCTCGTTCAGAATG
ACCCAGATTGTTTCTGCTTTGGGTGCAGCTGTTTAGTTCAGAGTTATATTACAGAGAATTATTTTCTGAG
ATAATCTTAAACTAGAATGTTCAAAACTAATTGATAATTGAAGTATCAAGATACGTAGAACACCTCAGAG
ATTTTTCTTCAGGAACTTCCACAAACTTTGAATCCTTGTATCTTTATTTGGTATTCATACTACTAGTAGC
AAAATACAGGTTTTTTGTTTTGTTTTGTTTTGGCTTCATAGAGTATCTCAAATTGAAACTTTTCTGCACA
AAGAATAAAATTAAGGATTTTATAAACTCAAATTGGCACCTACTGAATTAAAATACATAAAATCATTTAA
ATATAATTCAGCATATGGGAAGTAACATTGCACTAATATGGAAATCACTGCCAGAGACAGTCTATTTTCT
TTTAATTTGTTACTACTTAGTCACAAACCCCACATTATTCCAGTTTGGAATTACTTATTAAGGAGAATTG
GAAATACATATGCCCATGCTTAAATTTTATAGCTTTAATTTGTGTTATTTCTTTATTGACGGGAAGAGGT
ACATCTTTTTTTCCTTACTGAAAACCAAATATGGATTAATTGCCTCAAATTTGTATAAAGTGATTGGCTA
GTGATTCTTGTTTTCAGGAAGGGAGAGTGGTATAGATAGAAAATGACAAAGATGGCAATATACACTTAAT

GTTGTTATTGTATGTTGTTACTGAAGTACTTAGATTTTTAAAATTTCAAATCCTAAATCACTTCTTGTAG
GAGGGTTTTCATTAACTGCAGTATATACAGTTCACTACATATGGGTTGTTTGAGTTTTTTGTGTGCTGTA
TTTCTTTCTGTTTTTTTAATACCTGGTTTTGTACATATCTAACTCTGTTCTCTTTTGGTTGTTCAGAAAC
TGGATTTTTTTTTTTTCTTAAGCAGTGCTTAATTTGTGTTTTTAATTTTGATTCAGAAGTAGTCCCAGC
TCATAGGTGTTCATAACTGTTACATCCAGAACATTTGTCAGGCTCTCTGTCAGCTTTTCATGTACATATG
GTATAGAAACCATGGAGTTAGGCACTTCCCTGGAATTTTTTTTTTTATGAGAAAAATACTGTATTTAAAA
TGTAAAATAAACTTTTAAAAAGCAGGCACTAATATATATTTCTTCCAGCCTTTGATTACAAATTTGTCCT
TGCACATGTTAAGATGAATTATCTCCTAAAAATATCATTGTTCTTGGGAGCAGTGTATGTTACTTTACAT
AGCAGCGGTTCCTGTCATGTGTTCATGTTCAGAAATATTTTTGGTTTTAAACTTTCTTATTGCCTTTGGC
TGTTGATTAGTACAGTACAAGTTGCGATTTCAAAAAGATCTTGAAAGTAATATATTTAATCAATTAAAAT
GTTTATCTGTAAAAAAAAAAAAAAAAAA

Human GOCAP1 mRNA sequence - var3 (public gi: 15799258) (SEQ ID NO: 63)
GGAAGTCGATACGTGGCTGCCTTCTGTCCCCGCTGAGGAGGTGCAGCAGCCGGAGATGGCGGCGGTGCTG
AACGCAGAGCGACTCGAGGTGTCCGTCGACGGCCTCACGCTCAGCCCGGACCCGGAGGAGCGGCCTGGGG
CGGAGGGCGCCCCGCTGCTGCCGCCACCGCTGCCACCGCCCTCGCCACCTGGATCCGGTCGCGGCCCGGG
CGCCTCAGGGGAGCAGCCCGAGCCCGGGGAGGCGGCGGCTGGGGGCGCGGCGGAGGAGGCGCGGCGGCTG
GAGCAGCGCTGGGGTTTCGGCCTGGAGGAGTTGTACGGCCTGGCACTGCGCTTCTTCAAAGAAAAAGATG
GCAAAGCATTTCATCCAACTTATGAAGAAAAATTGAAGCTTGTGGCACTGCATAAGCAAGTTCTTATGGG
CCCATATAATCCAGACACTTGTCCTGAGGTTGGATTCTTTGATGTGTTGGGGAATGACAGGAGGAGAGAA
TGGGCAGCCCTGGGAAACATGTCTAAAGAGGATGCCATGGTGGAGTTTGTCAAGCTCTTAAATAGGTGTT
GCCATCTCTTTTCAACATATGTTGCGTCCCACAAAATAGAGAAGGAAGAGCAAGAAAAAAAAAAGGAAGGA
GGAAGAGGAGCGAAGGCGGCGTGAAGAGGAAGAAAGAGAACGTCTGCAAAAGGAGGAAGAGAAACGTAGG
AGAGAAGAAGAGGAAAGGCTTCGACGGGAGGAAGAGGAAAGGAGACGGATAGAAGAAGAAAGGCTTCGGT
TGGAGCAGCAAAAGCAGCAGATAATGGCAGCTTTAAACTCCCAGACTGCCGTGCAGTTCCAGCAGTATGC
AGCCCAACAGTATCCAGGGAACTACGAACAGCAGCAAATTCTCATCCGCCAGTTGCAGGAGCAACACTAT
CAGCAGTACATGCAGCAGTTGTATCAAGTCCAGCTTGCACAGCAACAGGCAGCATTACAGAAACAACAGG
AAGTAGTAGTGGCTGGGTCTTCCTTGCCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGAT
GTCAGTTAATGGACAGGCCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAACCAGAAGCTGCAGAA
GAAGCCCTGGAGAATGGACCAAAAGAATCTCTTCCAGTAATAGCAGCTCCATCCATGTGGACACGACCTC
AGATCAAAGACTTCAAAGAGAAGATTCAGCAGGATGCAGATTCCGTGATTACAGTGGGCCGAGGAGAAGT
GGTCACTGTTCGAGTACCCACCCATGAAGAAGGATCATATCTCTTTTGGGAATTTGCCACAGACAATTAT
GACATTGGGTTTGGGGTGTATTTTGAATGGACAGACTCTCCAAACACTGCTGTCAGCGTGCATGTCAGTG
AGTCCAGCGATGACGACGAGGAGGAAGAAGAAAACATCGGTTGTGAAGAGAAAGCCAAAAAGAATGCCAA
CAAGCCTTTGCTGGATGAGATTGTGCCTGTGTACCGACGGGACTGTCATGAGGAGGTGTATGCTGGCAGC
CATCAATATCCAGGGAGAGGAGTCTATCTCCTCAAGTTTGACAACTCCTACTCTTTGTGGCGGTCAAAAT
CAGTCTACTACAGAGTCTATTATACTAGATAAAAATGTTGTTACAAAGTCTGGAGTCTAGGGTTGGGCAG
AAGATGACATTTAATTTGGAAATTTCTTTTTACTTTTGTGGAGCATTAGAGTCACAGTTTACCTTATTGA
TATTGGTCTGATGGTTTGTGAACTCTTGCTGGGAATCAAAATTTCCTTGAGACTCTTTAGCATTCATACT
TTGGGGTTAAAGGAGATTCCTCAGACTCATCCAGCCCTTGGGTGCTGACCAGCAGAGTCACTAGTGGATG
CTGAAGTTACATGAGCTACATGTTAAATATTTAAAGTCTCCAAAATAAAACACCCCAACGTTGACCTTAC
CCGGCTGATGGTTAGCCCCTTGCTGCCTGCTCCATGTGTCTTATGAGAGCCCGTAGTTACAGTGTCCTCT
AATTTGAAATCCATAAGTTAACAAGTCTATATCAGGTGCATCTGGCTTTGATTAAAGGCCATTTTTAAAA
CTTAAAAACTCAACACCTCACAGATTATAATAGAAAAAGAAATGGCCTCAGTTTGATCTCGTTCAGAATG
ACCCAGATTGTTTCTGCTTTGGGTGCAGCTGTTTAGTTCAGAGTTATATTACAGAGAATTATTTTCTGAG
ATAATCTTAAACTAGAATGTTCAAAACTAATTGATAATTGAAGTATCAAGATACGTAGAACACCTCAGAG
ATTTTTCTTCAGGAACTTCCACAAACTTTGAATCCTTGTATCTTTATTTGGTATTCATACTACTAGTAGC
AAAATACAGGTTTTTTGTTTTGTTTTGTTTTGGCTTCATAGAGTATCTCAAATTGAAACTTTTCTGCACA
AAGAATAAAATTAAGGATTTTATAAACTCAAATTGGCACCTACTGAATTAAAATACATAAAATCATTTAA
ATATAATTCAGCATATGGGAAGTAACATTGCACTAATATGGAAATCACTGCCAGAGACAGTCTATTTTCT
TTTAATTTGTTACTACTTAGTCACAAACCCCACATTATTCCAGTTTGGAATTACTTATTAAGGAGAATTG
GAAATACATATGCCCATGCTTAAATTTTATAGCTTTAATTTGTGTTATTTCTTTATTGACGGGAAGAGGT
ACATCTTTTTTTCCTTACTGAAAACCAAATATGGATTAATTGCCTCAAATTTGTATAAAGTGATTGGCTA
GTGATTCTTGTTTTCAGGAAGGGAGAGTGGTATAGATAGAAAATGACAAAGATGGCAATATACACTTAAT
GTTGTTATTGTATGTTGTTACTGAAGTACTTAGATTTTTAAAATTTCAAATCCTAAATCACTTCTTGTAG
GAGGGTTTTCATTAACTGCAGTATATACAGTTCACTACATATGGGTTGTTTGAGTTTTTTGTGTGCTGTA
TTTCTTTCTGTTTTTTTAATACCTGGTTTTGTACATATCTAACTCTGTTCTCTTTTGGTTGTTCAGAAAC
TGGATTTTTTTTTTTTCTTAAGCAGTGCTTAATTTGTGTTTTTAATTTTGATTCAGAAGTAGTCCCAGC
TCATAGGTGTTCATAACTGTTACATCCAGAACATTTGTCAGGCTCTCTGTCAGCTTTTCATGTACATATG
GTATAGAAACCATGGAGTTAGGCACTTCCCTGGAATTTTTTTTTTTATGAGAAAAATACTGTATTTAAAA
TGTAAAATAAACTTTTAAAAAGCAGGCACTAATATATATTTCTTCCAGCCTTTGATTACAAATTTGTCCT
TGCACATGTTAAGATGAATTATCTCCTAAAAATATCATTGTTCTTGGGAGCAGTGTATGTTACTTTACAT
AGCAGCGGTTCCTGTCATGTGTTCATGTTCAGAAATATTTTTGGTTTTAAACTTTCTTATTGCCTTTGGC

TGTTGATTAGTACAGTACAAGTTGCGATTTCAAAAAGATCTTGAAAGTAATATATTTAATCAATTAAAAT
GTTTATCTGTAAAAAAAAAAAAAAAAAA

Human GOCAP1 mRNA sequence - var4 (public gi: 21961496) (SEQ ID NO: 64)
CGGACGCGTGGGTGCCATCTCTTTTCAACATATGTTGCGTCCCACAAAATAGAGAAGGAAGAGCAAGAAA
AAAAAAGGAAGGAGGAAGAGGAGCGAAGGCGGCGTGAAGAGGAAGAAAGAGAACGTCTGCAAAAGGAGGA
AGAGAAACGTAGGAGAGAAGAAGAGGAAAGGCTTCGACGGGAGGAAGAGGAAAGGAGACGGATAGAAGAA
GAAAGGCTTCGGTTGGAGCAGCAAAAGCAGCAGATAATGGCAGCTTTAAACTCCCAGACTGCCGTGCAGT
TCCAGCAGTATGCAGCCCAACAGTATCCAGGGAACTACGAACAGCAGCAAATTCTCATCCGCCAGTTGCA
GGAGCAACACTATCAGCAGTACATGCAGCAGTTGTATCAAGTCCAGCTTGCACAGCAACAGGCAGCATTA
CAGAAACAACAGGAAGTAGTAGTGGCTGGGTCTTCCTTGCCTACATCATCAAAAGTGAATGCAACTGTAC
CAAGTAATATGATGTCAGTTAATGGACAGGCCAAAACACACACTGACAGCTCCGAAAAAGAACTGGAACC
AGAAGCTGCAGAAGAAGCCCTGGAGAATGGACCAAAAGAATCTCTTCCAGTAATAGCAGCTCCATCCATG
TGGACACGACCTCAGATCAAAGACTTCAAAGAGAAGATTCAGCAGGATGCAGATTCCGTGATTACAGTGG
GCCGAGGAGAAGTGGTCACTGTTCGAGTACCCACCCATGAAGAAGGATCATATCTCTTTTGGGAATTTGC
CACAGACAATTATGACATTGGGTTTGGGGTGTATTTTGAATGGACAGACTCTCCAAACACTGCTGTCAGC
GTGCATGTCAGTGAGTCCAGCGATGACGACGAGGAGGAAGAAGAAAACATCGGTTGTGAAGAGAAAGCCA
AAAAGAATGCCAACAAGCCTTTGCTGGATGAGATTGTGCCTGTGTACCGACGGGACTGTCATGAGGAGGT
GTATGCTGGCAGCCATCAATATCCAGGGAGAGGAGTCTATCTCCTCAAGTTTGACAACTCCTACTCTTTG
TGGCGGTCAAAATCAGTCTACTACAGAGTCTATTATACTAGATAAAAATGTTGTTACAAAGTCTGGAGTC
TAGGGTTGGGCAGAAGATGACATTTAATTTGGAAATTTCTTTTTACTTTTGTGGAGCATTAGAGTCACAG
TTTACCTTATTGATATTGGTCTGATGGTTTGTGAACTCTTGCTGGGAATCAAAATTTCCTTGAGACTCTT
TAGCATTCATACTTTGGGGTTAAAGGAGATTCCTCAGACTCATCCAGCCCTTGGGTGCTGACCAGCAGAG
TCACTAGTGGATGCTGAAGTTACATGAGCTACATGTTAAATATTTAAAGTCTCCAAAATAAAACACCCCA
ACGTTGACCTTACCCGGCTGATGGTTAGCCCCTTGCTGCCTGCTCCATGTGTCTTATGAGAGCCCGTAGT
TACAGTGTCCTCTAATTTGAAATCCATAAGTTAACAAGTCTATATCAGGTGCAGCTGGCTTTGATTAAAG
GCCATTTTTAAAACTTAAAAACTCAACACCTCACAGATTATAATAGAAAAAGAAATGGCCTCAGTTTGAT
CTCGTTCAGAATGACCCAGATTGTTTCTGCTTTGGGTGCAGCTGTTTAGTTCAGAGTTATATTACAGAGA
ATTATTTTCTGAGATAATCTTAAACTAGAATGTTCAAAACTAATTGATAATTGAAGTATCAAGATACGTA
GAACACCTCAGAGATTTTTCTTCAGGAACTTCCACAAACTTTGAATCCTTGTATCTTTATTTGGTATTCA
TACTACTAGTAGCAAAATACAGGTTTTTTGTTTTGTTTTGTTTTGTTTTGGCTTCATAGAGTATCTCAAA
TTGAAACTTTTCTGCACAAAGAATAAAATTAAGGATTTTATAAACTCAAATTGGCACCTACTGAATTAAA
ATACATAAAATCATTTAAATATAATTCAGCATATGGGAAGTAACATTGCACTAATATGGAAATCACTGCC
AGAGACAGTCTATTTTCTTTTAATTTGTTACTACTTAGTCACAAACCCCCACATTATTCCAGTTTGGAATT
ACTTATTAAGGAGAATTGGAAATACATATGCCCATGCTTAAATTTTATAGCTTTAATTTGTGTTATTTCT
TTATTGACGGGAAGAGGTACATCTTTTTTTCCTTACTGAAAACAAATATGGATTAATTGCCTCAAATTTG
TATAAGTGATTGGCTAGTGATTCTTGTTTTCAGAAGGGAGAGTGGTATAGATAGAAAATGACAAAGATGG
CAATATACACTTAATGTTGTTATTGTATGTTGTTACTGAAGTACTTAGATTTTTAAAATTTCAAATCCTA
AATCACTTCTTGTAGGAGGGTTTCATTAACTGCAGTATATACAGTTCACTACATATGGGTTGTTTGAGT
TTTTTGTGTGCTGTATTTCTTTCTGTTTTTAATACCTGGTTTGTACATATCTAACTCTGTTCTCTTTT
GGTTGTTCAGAAACTGGATTTTTTTTTTTCTTAAGCAGTGCTTAATTTGTGTTTTTTAATTTTGATTCAG
AAGTAGTCCCAGCTCATAGGTGTTCATACTGTTACATCCAGAACATTTGTCAGGCTCTCTGTCAGCTTTC
ATGTACATATGGTATAGAAACCATGGAGTTAGGCACTTCCTGGATTTTTTTTTTATGAGAAAAATACTGT
ATTTAAAATGTAAAATAAACTTTTAAAAAGCAGGCACTAATATATATTTCTTCCAGCCTTTGATTACAAA
TTTGTCCTTGCACATGTTAAGATGAATTATCTCCTAAAAATATCATTGTTCTTGGGAGCAGTGTATGTTA
CTTTACATAGCAGCGGTTCCTGTCATGTGTTCATGTCAGAGTATTTTTGGTTTTAAACTTTCTTATTGCC
TTTGGCTGTTGATTAGTACAGTACAAGTGCGATTTCAAAAAGATCTTGAAAGTAATATATTTAATCAATT
AAAATGTTTATCTGTAAAAAAAAAAAAAAAAAAAAAAA

Human GOCAP1 mRNA sequence - var5 (public gi: 24496472) (SEQ ID NO: 65)
CCGCTGAGGAGGTGCAGCAGCCGGAGATGGCGGCGGTGCTGAACGCAGAGCGACTCGAGGTGTCCGTCGA
CGGCCTCACGCTCAGCCCGGACCCGGAGGAGCGGCCTGGGGCGGAGGGCGCCCCGCTGCTGCCGCCACCG
CTGCCACCGCCCTCGCCACCTGGATCCGGTCGCGGCCCGGGCGCCTCAGGGGAGCAGCCCGAGCCCGGGG
AGGCGGCGGCTGGGGGCGCGGCGGAGGAGGCGCGGCGGCTGGAGCAGCGCTGGGGTTTCGGCCTGGAGGA
GTTGTACGGCCTGGCACTGCGCCTCTTCAAAGAAAAAGATGGCAAAGCATTTCATCCAACTTATGAAGAA
AAATTGAAGCTTGTGGCACTGCATAAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCCTGAGG
TTGGATTCTTTGATGTGTTGGGGAATGACAGGAGGAGAGAATGGGCAGCCCTGGGAAACATGTCTAAAGA
GGATGCCATGGTGGAGTTTGTCAAGCTCTTAAATAGGTGTTGCCATCTCTTTTCAACATATGTTGCGTCC
CACAAAATAGAGAAGGAAGAGCAAGACAAAAAAAGGAAGGAGGAAGAGGAGCGAAGGCGGCGTGAAGAGG
AAGAAAGAGAGCGTCTGCAAAAGGAGGAAGAGAAACGTAGGAGAGAAGAAGAGGAAAGGCTTCGACGGGA
GGAAGAGGAAAGGATACGGATAGAAGAAGAAAGGCTTCGGTTGGAGCAGCAAAAGCAGCAGATAATGGCA
GCTTTAAACTCCCAGACTGCCGTGCAGTTCCAGCAGTATGCAGCCCAACGGTATCCAGGGAACTACGAAC

AGCAGCAAATTCTCATCCGCCAGTTGCAGGAGCAACACTATCAGCAGTACATGCAGCAGTTGTATCAAGT
CCAGCTTGCACAGCAACAGGCAGCATTACAGAAACAACAGGAAGTAGTAGTGGCTGGGTCTTCCTTGCCT
ACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGCCAGTTAATGGACAGGCCAAAACACACA
CTGACAGCTCCGAAAAAGAACTGGAACCAGAAGCTGCAGAAGAAGCCCTGGAGAATGGACCAAAAGAATC
TCTTCCAGTAATAGCAGCTCCATCCATGTGGACACGACCTCAGATCAAAGACTTTCAAAGAGAAGATTCA
GCAGGATGCAGATTCCGTGATTACAGTGGCCGAGGAGAAGTGGTCACTGTTCGAGTACCCACCCATGAAG
AAGGATCATATCTCTTTTGGGAATTTGCCACAGACAATTGTGACATTGGGTTTGGGGTGTATTTTGAATG
GACAGACTCTCCAAACACTGCTGTCAGCGTGCATGTCAGTGAGTCCAGCGATGACGACGAGGAGGAAGAA
GAAAACATCGGTTGTGAAGAGAAAGCCAAAAAGAATGCCAACAAGCCTTTGCTGGATGAGATTGTGCCTG
TGTACCGACGGGACTGTCATGAGGAGGTGTATGCTGGCAGCCATCAATATCCAGGGAGAGGAGTCTATCT
CCTCAAGTTTGACAACTCCTACTCTTTGTGGCGGTCAAAATCAGTCTACTACAGAGTCTATTATACTAGA
TAAAAATGTTGTTACAAAGTCTGGAGTCTAGGGTTGGGCAGAAGATGACATTTAATTTGGAAATTTCTTT
TTACTTTTGTGGAGCATTAGAGTCACAGTTTACCTTATTGATATTGGTCTGATGGTTTGTGAACTCTTGC
TGGGAATCAAAATTTCCTTGAGACTCTTTAGCATTCATACTTTGGGGTTAAAGGAGATTCCTCAGACTCA
TCCAGCCCTTGGGTGCTGACCAGCAGAGTCACTAGGGGATGCTGAAGTTACATGAGCTACATGTTAAATA
TTTAAAGTCTCCAAAATAAAACACCCCAACGTTGACCTTACCCGGCTGATGGTTAGCCCCTTGCTGCCTG
CTCCATGTGTCTTATGAGAGCCCGTAGTTACAGTGTCCTCTAATTTGAAATCCATAAGTTAACAAGTCTA
TATCAGGTGCAGCTGGCTTTGATTAAAGGCCATTTTTAAAACTTAAAAACTCAACACCTCACAGATTATA
ATAGAAAAAGAAATGGCCTCAGCTTGATCTCGTTCAGAATGACCCAGATTGTTTCTGCCTTGGGTGCAGC
TGTTTAGTTCAGAGTTATATTACAGAGAATTATTTCTGAGATAATCTTAAACTAGAATGTTCAAAACTA
ATTGATAATTGAAGTATCAAGATACGTAGAACACCTCAGAGATTTTTCTTCAGGAACTTCCACAAACTTT
GAATCCTTGTATCTTTATTTGGTATTCATACTACTAGTAGCAAAATACAGGTTTTTTGTTTTGTTTTGTT
TTGTTTTGGCTTCATAGAGTATCTCAAATTGAAACTTTTCTGCACAAAGAATAAAATTAAGGATTTTATA
AACTCGAATTGGCACCTACTGAATTAAAATACATAAAATCATTTAAATATAATTCAGCATATGGGAAGTA
ACATTGCACTAATATGGAAATCACTGCCAGAGACAGTCTATTTTCTTTTAATTTGTTACTACTTAGTCAC
AACCCCACATTATTCCAGTTTGGAATTACTTATTAAGGAGAATTGGAAATACATGTGCCCATGCTTAAAT
TTTATAGCTTTAATTTGTGTTATTTCTTTATTGACGGGAAGAGGTACATCTTTTTTTCCTTACTGAAAAC
AAATATGGATTAATTGCCTCAAATTTGTATAAAGTGATTGGCTAGTGATTCTTGTTTTCAGAGGGGAGAG
TGGTATAGATAGAAAATGACGAAGATGGCAATATACACTTAATGTTGTTATTGTATGTTGTTACTGAAGA
CTTAGATTTTTAAAATTTCAAATCCTAAATCACTTCTTGTAGGGGGGTTTTCATTAACTGCAGTATATAC
AGTTCACTACATATGGGTTGTTTGAGTTTTTTGTGTGCTGTATTTCTTTCTGTTTTTAATACCTGGTTT
TGTACATATCTAACTCTGTTCTCTTTTGGTTGTTCAGAAACTGGATTTTTTTCTTCTTAAGCAGTGCTTA
ATTTGTGTTTTTTAATTTTGATTCAGAAGTAGTCCCAGCTCATAGGCGTTCATACTGTTACATCCAGAAC
ATTTGTCAGGCTCTCTGTCAGCTTTCATGTACATATGGTATAGAAACCATGGAGTTAGGCACTTCCTGGA
TTTTTTTTATGAGAAAAATNCTGTATTTAAAATGTAAAATAAACTTTTAAAAAGCAGGCACTAATATATA
TTTCTTCCAGCCTTTGATTACAAATTTGTCCTTGCACATGTTAAGATGAATTATCTCCTAAAAATATCAT
TGTTCTTGGGAGCAGTGTATGTTACTTTACATAGCAGCGGTTCCTGTCATGTGTTCATGTCACGAATATT
TTTGGTTTTAAACTTTCTTATTGCCTTTGGCTGTTGATTAGTACAGTACAAGTGCGATTTCAAAAAGATC
TTGAAAGTAATATATTTAATCAATTAAAATGTTTATCTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

Human GOCAP1 mRNA sequence - var6 (public gi: 28374435) (SEQ ID NO: 66)
TCCGTCCCCGCTGAGGAGGTGCAGCAGCGGGAGATGGCGGCGGTGCTGAACGCAGAGCGACTCGAGGTGT
CCGTCGACGGCCTCACGCTCAGCCCGGACCCGGAGGAGCGGCCTGGGGCGGAGGGCGCCCCGCTGCTGCC
GCCACCGCTGCCACCGCCCTCGCCACCTGGATCCGGTCGCGGCCCGGGCGCCTCAGGGGAGCAGCCCGAG
CCCGGGGAGGCGGCGGCTGGGGGCGCGGCGGAGGAGGCGCGGCGGCTGGAGCAGCGCTGGGGTTTCGGCC
TGGAGGAGTTGTACGGCCTGGCACTGCGCTTCTTCAAAGAAAAAGATGGCAAAGCATTTCATCCAACTTA
TGAAGAAAAATTGAAGCTTGTGGCACTGCATAAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGT
CCTGAGGTTGGATTCTTTGATGTGTTGGGGAATGACAGGAGGAGAGAATGGGCAGCCCTGGGAAACATGT
CTAAAGAGGATGCCATGGTGGAGTTTGTCAAGCTCTTAAATAGGTGTTGCCATCTCTTTTCAACATATGT
TGCGTCCCACAAAATAGAGAAGGAAGAGCAAGAAAAAAAAAGGAAGGAGGAAGAGGAGCGAAGGCGGCGT
GAAGAGGAAGAAAGAGAACGTCTGCAAAAGGAGGAAGAGAAACGTAGGAGAGAAGAAGAGGAAAGGCTTC
GACGGGAGGAAGAGGAAAGGAGACGGATAGAAGAAGAAAGGCTTCGGTTGGAGCAGCAAAAGCAGCAGAT
AATGGCAGCTTTAAACTCCCAGACTGCCGTGCAGTTCCAGCAGTATGCAGCCCAACAGTATCCAGGGAAC
TACGAACAGCAGCAAATTCTCATCCGCCAGTTGCAGGAGCAACACTATCAGCAGTACATGCAGCAGTTGT
ATCAAGTCCAGCTTGCACAGCAACAGGCAGCATTACAGAAACAACAGGAAGTAGTAGTGGCTGGGTCTTC
CTTGCCTACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAGTTAATGGACAGGCCAAA
ACACACACTGACAGCTCCGAAAAAGAACTGGAACCGGAAGCTGCAGAAGAAGCCCTGGAGAATGGACCAA
AAGAATCTCTTCCAGTAATAGCAGCTCCATCCATGTGGACACGACCTCAGATCAAAGACTTCAAAGAGAA
GATTCAGCAGGATGCAGATTCCGTGATTACAGTGGCCGAGGAGAAGTGGTCACTGTTCGAGTACCCACC
CATGAAGAAGGATCATATCTCTTTTGGGAATTTGCCACAGACAATTATGACATTGGGTTTGGGGTGTATT
TTGAATGGACAGACTCTCCAAACACTGCTGTCAGCGTGCATGTCAGTGAGTCCAGCGATGACGACGAGGA
GGAAGAAGAAAACATCGGTTGTGAAGAGAAAGCCAAAAAGAATGCCAACAAGCCTTTGCTGGATGAGATT

GTGCCTGTGTACCGACGGGACTGTCATGAGGAGGTGTATGCTGGCAGCCATCAATATCCAGGGAGAGGAG
TCTATCTCCTCAAGTTTGACAACTCCTACTCTTTGTGGCGGTCAAAATCAGTCTACTACAGAGTCTATTA
TACTAGATAAAAATGTTGTTACAAAGTCTGGAGTCTAGGGTTGGCAGAAGATGACATTTAATTTGGAAA
TTTCTTTTTACTTTTGTGGAGCATTAGAGTCACAGTTTACCTTATTGATATTGGTCTGATGGTTTGTGAA
CTCTTGCTGGGAATCAAAATTTCCTTGAGACTCTTTAGCATTCATACTTTGGGGTTAAAGGAGATTCCTC
AGACTCATCCAGCCCTTGGGTGCTGACCAGCAGAGTCACTAGTGGATGCTGAAGTTACATGAGCTACATG
TTAAATATTTAAAGTCTCCAAAATAAAACACCCCAACGTTGACCTTACCCGGCTGATGGTTAGCCCCTTG
CTGCCTGCTCCATGTGTCTTATGAGAGCCCGTAGTTACAGTGTCCTCTAATTTGAAATCCATAAGTTAAC
AAGTCTATATCAGGTGCAGCTGGCTTTGATTAAAGGCCATTTTTAAAACTTAAAAACTCAACACCTCACA
GATTATAATAGAAAAAGAAATGGCCTCAGTTTGATCTCGTTCAGAATGACCCAGATTGTTTCTGCTTTGG
GTGCAGCTGTTTAGTTCAGAGTTATATTACAGAGAATTATTTTCTGAGATAATCTTAAACTAGAATGTTC.
AAAACTAATTGATAATTGAAGTATCAAGATACGTAGAACACCTCAGAGATTTTTCTTCAGGAACTTCCAC
AAACTTTGAATCCTTGTATCTTTATTTGGTATTCATACTACTAGTAGCAAAATACAGGTTTTTTGTTTTG
TTTTGTTTTGGCTTCATAGAGTATCTCAAATTGAAACTTTTCTGCACAAAGAATAAAATTAAGGATTTTA
TAAACTCAAATTGGCACCTACTGAATTAAAATACATAAAATCATTTAAATATAATTCAGCATATGGGAAG
TAACATTGCACTAATATGGAAATCACTGCCAGAGACAGTCTATTTTCTTTTAATTTGTTACTACTTAGTC
ACAAACCCCACATTATTCCAGTTTGGAATTACTTATTAAGGAGAATTGGAAATACATATGCCCATGCTTA
AATTTTATAGCTTTAATTTGTGTTATTTCTTTATTGACGGGAAGAGGTACATCTTTTTTTCCTTACTGAA
AACAAATATGGATTAATTGCCTCAAATTTGTATAAGTGATTGGCTAGTGATTCTTGTTTTCAGAAGGGAG
AGTGGTATAGATAGAAAATGACAAAGATGGCAATATACACTTAATGTTGTTATTGTATGTTGTTACTGAA
GTACTTAGATTTTTAAAATTTCAAATCCTAAATCACTTCTTGTAGGAGGGTTTTCATTAACTGCAGTATA
TACAGTTCACTACATATGGGTTGTTTGAGTTTTTTGTGTGCTGTATTTCTTTCTGTTTTTTAATACCTGG
TTTTGTACATATCTAACTCTGTTCTCTTTTGGTTGTTCAGAAACTGGATTTTTTTTTCTTAAGCAGTGCT
TAATTTGTGTTTTTTAATTTTGATTCAGAAGTAGTCCCAGCTCATAGGTGTTCATACTGTTACATCCAGA
ACATTTGTCAGGCTCTCTGTCAGCTTTCATGTACATATGGTATAGAAACCATGGAGTTAGGCACTTCCTG
GATTTTTTTTTTTATGAGAAAAATACTGTATTTAAAATGTAAAATAAACTTTTAAAAAGCAGGCACTAAT
ATATATTTCTTCCAGCCTTTGATTACAAATTTGTCCTTGCACATGTTAAGATGAATTATCTCCTAAAAAT
ATCATTGTTCTTGGGAGCAGTGTATGTTACTTTACATAGCAGCGGTTCCTGTCATGTGTTCATGTCAGAA
TATTTTTGGTTTTAAACTTTCTTATTGCCTTTGGCTGTTGATTAGTACAGTACAAGTGCGATTTCAAAAA
GATCTTGAAAGTAATATATTTAATCAATTAAAATGTTTATCTGTCAAAAAAAAAAAAAAA

Human GOCAP1 mRNA sequence - var7 (public gi: 25058702) (SEQ ID NO: 67)
CGCTGAGGAGGTGCAGCAGCCGGAGATGGCGGCGGTGCTGAACGCAGAGCGACTCGAGGTGTCCGTCGAC
GGCCTCACGCTCAGCCCGGACCCGGAGGAGCGGCCTGGGGCGGAGGGCGCCCCGCTGCTGCCGCCACCGC
TGCCACCGCCCTCGCCACCTGGATCCGGTCGCGGCCCGGGCGCCTCAGGGGAGCAGCCCGAGCCCGGGGA
GGCGGCGGCTGGGGGCGCGGCGGAGGAGGCGCGGCGGCTGGAGCAGCGCTGGGGTTTCGGCCTGGAGGAG
TTGTACGGCCTGGCACTGCGCTTCTTCAAAGAAAAAGATGGCAAAGCATTTCATCCAACTTATGAAGAAA
AATTGAAGCTTGTGGCACTGCATAAGCAAGTTCTTATGGGCCCATATAATCCAGACACTTGTCCTGAGGT
TGGATTCTTTGATGTGTTGGGGAATGACAGGAGGAGAGAATGGGCAGCCCTGGGAAACATGTCTAAAGAG
GATGCCATGGTGGAGTTTGTCAAGCTCTTAAATAGGTGTTGCCATCTCTTTTCAACATATGTTGCGTCCC
ACAAAATAGAGAAGGAAGAGCAAGAAAAAAAAAAGGAAGGAGGAAGAGGAGCGAAGGCGGCGTGAAGAGG
AAGAAAGAGAACGTCTGCAAAAGGAGGAAGAGAAACGTAGGAGAGAAGAAGAGGAAAGGCTTCGACGGGA
GGAAGAGGAAAGGAGACGGATAGAAGAAGAAAGGCTTCGGTTGGAGCAGCAAAAGCAGCAGATAATGGCA
GCTTTAAACTCCCAGACTGCCGTGCAGTTCCAGCAGTATGCAGCCCAACAGTATCCAGGGAACTACGAAC
AGCAGCAAATTCTCATCCGCCAGTTGCAGGAGCAACACTATCAGCAGTACATGCAGCAGTTGTATCAAGT
CCAGCTTGCACAGCAACAGGCAGCATTACAGAAACAACAGGAAGTAGTAGTGGCTGGGTCTTCCTTGCCT
ACATCATCAAAAGTGAATGCAACTGTACCAAGTAATATGATGTCAGTTAATGGACAGGCCAAAACACACA
CTGACAGCTCCGAAAAAGAACTGGAACCAGAAGCTGCAGAAGAAGCCCTGGAGAATGGACCAAAAGAATC
TCTTCCAGTAATAGCAGCTCCATCCATGTGGACACGACCTCAGATCAAAGACTTCAAAGAGAAGATTCAG
CAGGATGCAGATTCCGTGATTACAGTGGGCCGAGGAGAAGTGGTCACTGTTCGAGTACCCACCCATGAAG
AAGGATCATATCTCTTTTGGGAATTTGCCACAGACAATTATGACATTGGGTTTGGGGTGTATTTTGAATG
GACAGACTCTCCAAACACTGCTGTCAGCGTGCATGTCAGTGAGTCCAGCGATGACGACGAGGAGGAAGAA
GAAAACATCGGTTGTGAAGAGAAAGCCAAAAAGAATGCCAACAAGCCTTTGCTGGATGAGATTGTGCCTG
TGTACCGACGGGACTGTCATGAGGAGGTGTATGCTGGCAGCCATCAATATCCAGGGAGAGGAGTCTATCT
CCTCAAGTTTGACAACTCCTACTCTTTGTGGCGGTCAAAATCAGTCTACTACAGAGTCTATTATACTAGA
TAAAAATGTTGTTACAAAGTCTGGAGTCTAGGGTTGGCAGAAGATGACATTTAATTTGGAAATTTCTTT
TTACTTTTGTGGAGCATTAGAGTCACAGTTTACCTTATTGATATTGGTCTGATGGTTTGTGAACTCTTGC
TGGGAATCAAAATTTCCTTGAGACTCTTTAGCATTCATACTTTGGGGTTAAAGGAGATTCCTCAGACTCA
TCCAGCCCTTGGGTGCTGACCAGCAGAGTCACTAGTGGATGCTGAAGTTACATGAGCTACATGTTAAATA
TTTAAAGTCTCCAAAATAAAACACCCCAACGTTGACCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human GOCAP1 mRNA sequence - var8 (public gi: 2738926) (SEQ ID NO: 68)
GAATTCCGTTGCTGTCGGAGCCCGTAGTTACAGTGTCCTCTAATTTGAAATCCATAAGTTACCAAGTCTA
TATCAGGTACAGCTGGCTTTCATTAAAGGCCATTTTTAAAACTTCAAAAACTCAACACCTCACAGATTAT
AATAGAAAAAGAAATGGCCTCAGTTTGATCTCGTTCAGAATGACCCAGATTGTTTTCTGCTTTGGGTGCA
GCTGTTTAGTTCAGAGTTATATTACAGAGAATTATTTTCTGAGAAATCTTAAAACTAGAATGTTCAAAAC
TAATTCGATAATTGAAGTATCAAGATACGTAGAACACCTCAGAGATTTTTCTTCAGGAACTTCCACAAAC
TTTAGAATCCTTGTATCTTTATTTGGTATTCATACTACTAGTCGCAAAATACAGGTTTTTTGTTTTGTTT
TGTTTTGTTTTGGCTTCATAGAGTATCTCAAATTGAAACTTTTCTGCCCAAAGAATAAAATTAAGGATTT
TATAAAACTCAAATTGGCACCTACTGAATTAAAATACATAAAATGCATTAAATATAATTCAGCATATGGC
AGTAACATTGCACTAATATGGAAATCACTGCCAGAGACAGTCTATTTTCTTTTAATTTGTTACTACTTAG
TCACAACCCCACATTATTCCAGTTTGGAATTACTTATTAAGGAGAATTGGAAATACATATGCCCATGCTT
AAATTTTATAGCTTTAATTTGTGTTATTTCTTTATTGACGGAAGAGGTACATCTTTTTTTCCTTACTCA
AAACAAATATGGATTAATTGCCTCAAATTTGTATAAGTGATTGGCTAGTGATTCTTGTTTTCAGAGGGAG
AGTGGTATAGATAGAAAATGACAAAGATGGCAATATACACTTAATGTTGTTATTGTATGTTGTTACTGAA
GTACTTAGATTTTTAAAATTTCAAATCCTAAATCACTTCTTGTAGGAGGGTTTTCATTAACTGCAGATAT
ACAGTTCACTACATATGGGTTGTTTGAGTTTTTTGTGTGCTGTATTTCTTTCTGTTTTTTAATACCTGGT
TTTGTACATATCTAACTCTGTTCTCTTTTGGTTGTTCAGAAACTGGATTTTTTTTTTCTTAAGCAGTGCT
TAATTTGTGTTTTTTAATTTTGATTCAGAAGTAGTCCCAGCTCATAGGTGTTCATACTGTTACATCCAGA
ACATTTGTCAGGCTCTCTGTCAGCTTTCATGTACATATGGTATAGAAACCATGGAGTTAGGCACTTCCTG
GATTTTTTTTTTATGAGAAAAATACTGTATTTAAAATGTAAAATAAACTTTTAAAAAGC

Human GOCAP1 Protein sequence - var1 (public gi: 24496473) (SEQ ID NO: 240)
MAAVLNAERLEVSVDGLTLSPDPEERPGAEGAPLLPPPLPPPSPPGSGRGPGASGEQPEPGEAAAGGAAE
EARRLEQRWGFGLEELYGLALRLFKEKDGKAFHPTYEEKLKLVALHKQVLMGPYNPDTCPEVGFFDVLGN
DRRREWAALGNMSKEDAMVEFVKLLNRCCHLFSTYVASHKIEKEEQDKKRKEEEERRRREEEERERLQKE
EEKRRREEEERLRREEEERIRIEEERLRLEQQKQQIMAALNSQTAVQFQQYAAQRYPGNYEQQQILIRQL
QEQHYQQYMQQLYQVQLAQQQAALQKQQEVVVAGSSLPTSSKVNATVPSNMMPVNGQAKTHTDSSEKELE
PEAAEEALENGPKESLPVIAAPSMWTRPQIKDFQREDSAGCRFRDYSGRGEVVTVRVPTHEEGSYLFWEF
ATDNCDIGFGVYFEWTDSPNTAVSVHVSESSDDDEEEEENIGCEEKAKKNANKPLLDEIVPVYRRDCHEE
VYAGSHQYPGRGVYLLKFDNSYSLWRSKSVYYRVYYTR

Human GOCAP1 Protein sequence - var2 (public gi: 21961497) (SEQ ID NO: 241)
RTRGCHLFSTYVASHKIEKEEQEKKRKEEEERRRREEEERERLQKEEEKRRREEEERLRREEEERRRIEE
ERLRLEQQKQQIMAALNSQTAVQFQQYAAQQYPGNYEQQQILIRQLQEQHYQQYMQQLYQVQLAQQQAAL
QKQQEVVVAGSSLPTSSKVNATVPSNMMSVNGQAKTHTDSSEKELEPEAAEEALENGPKESLPVIAAPSM
WTRPQIKDFKEKIQQDADSVITVGRGEVVTVRVPTHEEGSYLFWEFATDNYDIGFGVYFEWTDSPNTAVS
VHVSESSDDDEEEEENIGCEEKAKKNANKPLLDEIVPVYRRDCHEEVYAGSHQYPGRGVYLLKFDNSYSL
WRSKSVYYRVYYTR

Human GOCAP1 Protein sequence - var3 (public gi: 15799259) (SEQ ID NO: 242)
MAAVLNAERLEVSVDGLTLSPDPEERPGAEGAPLLPPPLPPPSPPGSGRGPGASGEQPEPGEAAAGGAAE
EARRLEQRWGFGLEELYGLALRFFKEKDGKAFHPTYEEKLKLVALHKQVLMGPYNPDTCPEVGFFDVLGN
DRRREWAALGNMSKEDAMVEFVKLLNRCCHLFSTYVASHKIEKEEQEKKRKEEEERRRREEEERERLQKE
EEKRRREEEERLRREEEERRRIEEERLRLEQQKQQIMAALNSQTAVQFQQYAAQQYPGNYEQQQILIRQL
QEQHYQQYMQQLYQVQLAQQQAALQKQQEVVVAGSSLPTSSKVNATVPSNMMSVNGQAKTHTDSSEKELE
PEAAEEALENGPKESLPVIAAPSMWTRPQIKDFKEKIQQDADSVITVGRGEVVTVRVPTHEEGSYLFWEF
ATDNYDIGFGVYFEWTDSPNTAVSVHVSESSDDDEEEEENIGCEEKAKKNANKPLLDEIVPVYRRDCHEE
VYAGSHQYPGRGVYLLKFDNSYSLWRSKSVYYRVYYTR

Human GOCAP1 Protein sequence - var4 (public gi: 10438061) (SEQ ID NO: 243)
MAAVLNAERLEVSVDGLTLSPDPEERPGAEGAPLLPPPLPPPSPPGSGRGPGASGEQPEPGEAAAGGAAE
EARRLEQRWGFGLEELYGLALRFFKEKDGKAFHPTYEEKLKLVALHKQVLMGPYNPDTCPEVGFFDVLGN
DRRREWAALGNMSKEDAMVEFVKLLNRCCHLFSTYVASHKIEKEEQDKKRKEEEERRRREEEERERLQKE
EEKRRREEEERLRREEEERRRIEEERLRLEQQKQQIMAALNSQTAVQFQQYAAQQYPGNYEQQQILIRQL
QEQHYQQYMQQLYQVQLAQQQAALQKQQEVVVAGSSLPTSSKVNATVPSNMMSVNGQAKTHTDSSEKELE
PEAAEEALENGPKESLPVIAAPSMWTRPQIKDFKEKIQQDADSVITVGRGEVVTVRVPTHEEGSYLFWEF
ATDNYDIGFGVYFEWTDSPNTAVSVHVSESSDDDEEEEENIGCEEKAKKNANKPLLDEIVPVYRRDCHEE
VYAGSHQYPGRGVYLLKFDNSYSLWRSKSVYYRVYYTR

Unigene Name: GOSR2 **Unigene ID:** Hs.432552

Human GOSR2 mRNA sequence - var1 (public gi: 2316087) (SEQ ID NO: 69)
ATGGATCCCCTGTTCCAGCAAACGCACAAGCAGGTCCACGAGATCCAGTCTTGCATGGACGCCTGGAGA
CGGCAGACAAGCAGTCTGTGCACATAGTAGAAAACGAAATCCAAGCAAGCATAGACCAGATATTCAGCCG
TCTAGAACGTCTGGAGATTTTGTCCAGCAAGGAGCCCCCTAACAAAAGGCAAAATGCCAGACTTCGGGTT
GACCAGTTAAAGTATGATGTCCAGCACCTGCAGACTGCGCTCAGAAACTTCCAGCATCGGCGCCATGCAA
GGGAGCAGCAGGAGAGACAGCGAGAAGAGCTTCTGTGTCGAACGTTCACCACTAACGGCTCTGACACCAC
CATACCAATGGACGAATCACTGCAGTTTAACTCCTCCCTCCAGAAAGTTCACAACGGCATGGATGACCTC
ATTTTAGATGGGCACAATATTTTAGATGGACTGAGGACCCAGAGACTGACCTTGAAGGGGACTCAGAAGA
AGATCCCTGACATTGCCAACATGCTGGGCTTGTCCAACACAGTGATGCGGCTCATCGAGAAGCGGGCTTT
CCAGGACAAGTACTTTATGATAGGTGGGATGCTGCTGACCTGTGTGGTCATGTTCCTCGTGGTGCAGTAC

Human GOSR2 mRNA sequence - var2 (public gi: 3483524) (SEQ ID NO: 70)
TTTTTTTTTTCAGGACAGATTGGCCTTTATACTAAATTCCACAATATACCTGGTATTAGTACAGCCTGAA
TCCGGGGCTGGTCACAGAAGGAAAAAGGTTGTAGTCCCTGAAAACAGAGTGTTACAAGGACATACACACT
ACAGATGTCTCCACGGTGGGATCTGCCCACACTGGCTGGGCAAAATGAGGGCCTGGCTGGCAGGTGCTAA
TATATTTCAGGGAAGAGAAGGGAACCAAAGAATTAGAGATACTAAAACTAGAGCTGAGACTGTAATTGGA
AAATCACAAATCTTTGCCTACAGCTACTTTCTAAGGGGCAAAGGCCCACAAAGCCTGGGCGCAGGTGCCA
AGCCACAGTCTCTGAACCTTAAAAGCCAACCACTCTATTAACAACTAGAAAAAATCAGTGACTAGGGTCA
AGACTGAACACTCCCGGGAAATAACACTGGCCTCACTTTAGAAAAGAGAAACACCCAGCTGTAGTGTGGA
AAATCTTACTTGTTATCGGCAATAGCACTACATCTTGTTTCCCCTAGGTAGCTGCTTCCAGGGAATGGTG
ACAAGTATTGGCAGTCAGTCATCTACATGTCACTGAGGCACAGGGGAGGGTGGCCAGGAGCACGAGGATG
TGAATCGACCTACTATTTAATATAATGGCTGTGAGAAAAGGCCTCTTTCCTTTCCTTTCCACTTTTGCTC
CACCCTATCAGGAG

Human GOSR2 mRNA sequence - var3 (public gi: 21961348) (SEQ ID NO: 71)
GGCCTGCCGGGCCGGCGACATGGATCCCCTGTTCCAGCAAACGCACAAGCAGGTCCACGAGATCCAGTCT
TGCATGGGACGCCTGGAGACGGCAGACAAGCAGTCTGTGCACATAGTAGAAAACGAAATCCAAGCAAGCA
TAGACCAGATATTCAGCCGTCTAGAACGTCTGGAGATTTTGTCCAGCAAGGAGCCCCCTAACAAAAGGCA
AAATGCCAGACTTCGGGTTGACCAGTTAAAGTATGATGTCCAGCACCTGCAGACTGCGCTCAGAAACTTC
CAGCATCGGCGCCATGCAAGGGAGCAGCAGGAGAGACAGCGAGAAGAGCTTCTGTCTCGAACCTTCACCA
CTAACGACTCTGACACCACCATACCAATGGACGAATCACTGCAGTTTAACTCCTCCCTCCAGAAAGTTCA
CAACGGCATGGATGACCTCATTTTAGATGGGCACAATATTTTAGATGGACTGAGGACCCAGAGACTGACC
TTGAAGGTGGGGTCCCTGCTGGGGGACAGAGAGAAGGCCTCTTGTTTTAGCCTCATCCAACAGTTTAGTA
ACTGTGTTTATATTTTGATTACGTGTCCTCAAATTGTGATATTTTGATGACAAGACAGAGCCCTTGAGTT
TGGGATCCTTTCTGTTGGAGTTGAGTTATTGTGAGCCTGAAAGTACCCAGTTCCTTTGCCAGTGCTTGAA
ACAAACCATGAAGTGGCCTCTCTTAGGATCCAGGTCTTTTCCCATTTACTGAACTTATCATGAAAGTGAG
TGCTACTACGAGGGGTCCAATCACAGGCTGAGAAATTGTGTTACAGAATCTACTCTTGGAAGAATGAAGA
CGTGGCTGTCCTTTGGTACCTCGCTTTAAGGTGGCTTTCCCTTAGGACCCCTACTGTGGACTGCCTTATA
ACTAAAACCTTTTGTATTTTAGTAACTGAATCCCCACTGTGCAGTGTTAGGGCTGCCTGGTTGTTTGCAG
TAGATTAGAGCTTTAGAAGCTTCTAGAGCTTCTAAAGCCCGTGCTGGTGATCCCAGCGACTCTTCACTCC
CTAGCCTTAGGTATTCCTAGAAGCCCTGACCAGTTGGCACTGCTGAGACTCCAGCCCCTGGGAGTGGTTT
ACAGAAACATTACACAGACTCTGATGTCAGTCATGATGTTTCAGCCTCTGCCCTTTTCCTGTATCAACCC
TGATGGATAATAGGGCGTGGGTTCTGTCTGTTATCAGGGTGTGGTCCCCTGTGAATGAAGCACTCCCAGC
CACTGAGCTGTGAGAAACAGTCACTCGGAAGTGTGAGCTTTATCTTAGTTTTTGTTGGATCATGTTGAGT
CTGTCAGCTCCACAGGACTTCAGTACGTTTCTGAACAGTCCCTGCCATCTCTACGGGGGAGAGGGTCAGG
CAAGCTGCAAGTGACACTCACCTCCTGCTGACAGTTGCAGTGTCTCAGATGGCCTGGAAGGGTGGTCTCC
AGCAGCCTGCTGGGCGCTCCCCTTTCATGAGAGCCACCTGCAGTGACCTGAACTGATACATGTTGATTAG
TCTGCCCTTTCTTTAGAAAACTGCTACTCTCCTTTTCATATCTCAGAAAAACAGTAGAGGCCTTTTAGGA
CCAAACTCCATGTCACACTGATGAAGAGCCAGTGGGGGTTAGAGCGTCCTGTTAAGGCACATGCTAGCTT
CCCACTCAAGTCTGGCAGCGCTGGGGCATCAGCACACCTCTTGCCACCCACACTGATACCAGAGGGGAAG
GCTGTGAGGTGGCTGGGGGTTGAGACTTGAGGTTTCTAACTTTCCTCTGCACACCTGTGGCTACCTGGTG
TTTGTCTCTTGATTCCCTCCACCTGCCTCACACCCTGCCTCCGTCGGGATTTTCCACCTACACCATTCAA
AAGGAACATAGGAGAGGGCATGAAGGGGCTAGGCTGAAGCACTCTGATGACTGGGGCCAATTTGTGGCTG
AAAATGAATACATTTTTTGAAATTTATGGTCATTTTCAAGTGATTTAGAAGGTTGATCCTTAGCCTCATA
CAGTGATGAAATAATCTGTGTGTTCAGAGCCAAGCAGGACTTTAGCAAGAGTCTGATTGTATTGTCACTA
TCTCGGGGAAAAAAAAATACAAATACATTTCTCTGATCTCTGATGGCAATGAAGTTTGACTTGTAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human GOSR2 mRNA sequence - var4 (public gi: 16905519) (SEQ ID NO: 72)
GTTCCGAGGAAGCCAGAGCCGGAGCCGTGGCCTGCCGGGCCGGCGACATGGATCCCCTGTTCCAGCAAAC
GCACAAGCAGGTCCACGAGATCCAGTCTTGCATGGGACGCCTGGAGACGGCAGACAAGCAGTCTGTGCAC
ATAGTAGAAAACGAAATCCAAGCAAGCATAGACCAGATATTCAGCCGTCTAGAACGTCTGGAGATTTTGT
CCAGCAAGGAGCCCCCTAACAAAAGGCAAAATGCCAGACTTCGGGTTGACCAGTTAAAGTATGATGTCCA
GCACCTGCAGACTGCGCTCAGAAACTTCCAGCATCGGCGCCATGCAAGGGAGCAGCAGGAGAGACAGCGA
GAAGAGCTTCTGTCTCGAACCTTCACCACTAACGACTCTGACACCACCATACCAATGGACGAATCACTGC
AGTTTAACTCCTCCCTCCAGAAAGTTCACAACGGCATGGATGACCTCATTTTAGATGGGCACAATATTTT
AGATGGACTGAGGACCCAGAGACTGACCTTGAAGGGGACTCAGAAGAAGATCCTTGACATTGCCAACATG
CTGGGCTTGTCCAACACAGTGATGCGGCTCATCGAGAAGCGGGCTTTCCAGGACAAGTACTTTATGATAG
GCACCCAAGGATCCTGCCAGACAGCACACTTTGGAGGAAGGTCTGCAGGGAGCAGCTGAGCCATTTGTTC
TTGAACTCTGGGAGGCAGAAGTCCCCGCACCCATCATGCGTGGACTGATAGGACATCTTTTCGTGGTGTG
CACCAGTGCTTTCCACACTTGACAGTGGTTGGCTTTGATGAACCCTCATGCTGCACCTTCAGAGCCAGTC
CTCTAGTTTGGAATAAAAATTGCAGAGGTGGAAAAAAAAAAAAAAAAAAA

Human GOSR2 mRNA sequence - var5 (public gi: 12711466) (SEQ ID NO: 73)
AGCCGGAGCCGTGGCCTGCGGGGCCGGCGACATGGATCCCCTGTTCCAGCAAACGCACAAGCAGGTCCAC
GAGATCCAGTCTTGCATGGGACGCCTGGAGACGGCAGACAAGCAGTCTGTGCACATAGTAGAAAACGAAA
TCCAAGCAAGCATAGACCAGATATTCAGCCGTCTAGAACGTCTGGAGATTTTGTCCAGCAAGGAGCCCCC
TAACAAAAGGCAAAATGCCAGACTTCGGGTTGACCAGTTAAAGTATGATGTCCAGCACCTGCAGACTGCG
CTCAGAAACTTCCAGCATCGGCGCCATGCAAGGGAGCAGCAGGAGAGACAGCGAGAAGAGCTTCTGTCTC
GAACCTTCACCACTAACGACTCTGACACCACCATACCAATGGACGAATCACTGCAGTTTAACTCCTCCCT
CCAGAAAGTTCACAACGGCATGGATGACCTCATTTTAGATGGGCACAATATTTTAGATGGACTGAGGACC
CAGAGACTGACCTTGAAGGGGACTCAGAAGAAGATCCTTGACATTGCCAACATGCTGGGCTTGTCCAACA
CAGTGATGCGGCTCATCGAGAAGCGGGCTTTCCAGGACAAGTACTTTATGATAGGCACCCAAGGATCCTG
CCAGACAGCACACTTTGGAGGAAGGTCTGCAGGGAGCAGCTGAGCCATTTGTTCTTGAACTCTGGGAGGC
AGAAGTCCCCGCACCCATCATGCGTGGACTGATAGGACATCTTTTCGTGGTGTGCACCAGTGCTTTCCAC
ACTTGACAGTGGTTGGCTTTGATGAACCCTCATGCTGCACCTTCAGAGCCAGTCCTCTAGTTTGGAATAA
AAATTGCAGAGGTGGAAAAAAAAAAAAAAAAAAA

Human GOSR2 mRNA sequence - var6 (public gi: 37805253) (SEQ ID NO: 74)
CAATAGAGACAAGGTCTTGCTCTGTCACCCAGGTTGGAGTACAGTGGCATGATCTTGATTCACTACAACC
TCTACCTCTTGGGTTCAAGCGATCCTCCCACCTCGGTCTTCTGAGTAGCTGGGAATACAGTTATAATTAT
TCAATATGTTCCCACTGACTGAGGAAAACAAGCATGTGGCCCAGTTGTTGCTCAATACTGGTACTTGTCC
AAGATGTATCTTCAGATTCTGTGGTGTGGATTTTCATGCACCTTACAAACTTCCATACAAGATGAAGAAA
CTGAGATACAGAGAGGTTAAGCAACCTCCCAAAGTTCTAGGGTTACAGGTGTTAGCCACTGTACCTGGCC
TCTAAGGTGATTCTGATGTGTGTATTTTGGAACCACTGTCTCCTAGACAGAAAGCTTCTGTCTCAAAGAT
GATCACATTGGTGTAAAGAGCAAAACTTGTTAAGTCCAAAATAAATTCTTACTGTTTATATCCTAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human GOSR2 mRNA sequence - var7 (public gi: 16905521) (SEQ ID NO: 75)
GTTCCGAGGAAGCCAGAGCCGGAGCCGTGGCCTGCCGGGCCGGCGACATGGATCCCCTGTTCCAGCAAAC
GCACAAGCAGGTCCACGAGATCCAGTCTTGCATGGGACGCCTGGAGACGGCAGACAAGCAGTCTGTGCAC
ATAGTAGAAAACGAAATCCAAGCAAGCATAGACCAGATATTCAGCCGTCTAGAACGTCTGGAGATTTTGT
CCAGCAAGGAGCCCCCTAACAAAAGGCAAAATGCCAGACTTCGGGTTGACCAGTTAAAGTATGATGTCCA
GCACCTGCAGACTGCGCTCAGAAACTTCCAGCATCGGCGCCATGCAAGGGAGCAGCAGGAGAGACAGCGA
GAAGAGCTTCTGTCTCGAACCTTCACCACTAACGACTCTGACACCACCATACCAATGGACGAATCACTGC
AGTTTAACTCCTCCCTCCAGAAAGTTCACAACGGCATGGATGACCTCATTTTAGATGGGCACAATATTTT
AGATGGACTGAGGACCCAGAGACTGACCTTGAAGGGGACTCAGAAGAAGATCCTTGACATTGCCAACATG
CTGGGCTTGTCCAACACAGTGATGCGGCTCATCGAGAAGCGGGCTTTCCAGGACAAGTACTTTATGATAG
GTGGGATGCTGCTGACCTGTGTGGTCATGTTCCTCGTGGTGCAGTACCTGACATGAGCCAGCCACGCTCA
GTGGCTGAACAGCATTCCCACAGCCTGCAAGTGTGTGTGTGTGTGAAAGAGAGAGGGGGGCCCAGAGGCC
GCCTTTTGAAATGTTTGCCTGTCTGAACTGTGAAGACACTTGGGAGTGATTGTGGTCTAATTTCCAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human GOSR2 protein sequence - var1 (public gi: 16307241) (SEQ ID NO: 244)
MDPLFQQTHKQVHEIQSCMGRLETADKQSVHIVENEIQASIDQIFSRLERLEILSSKEPPNKRQNARLRV
DQLKYDVQHLQTALRNFQHRRHAREQQERQREELLSRTFTTNDSDTTIPMDESLQFNSSLQKVHNGMDDL
ILDGHNILDGLRTQRLTLKGTQKKILDIANMLGLSNTVMRLIEKRAFQDKYFMIGGMLLTCVVMFLVVQY
LT

Human GOSR2 protein sequence - var2 (public gi: 16905522) (SEQ ID NO: 245)
MDPLFQQTHKQVHEIQSCMGRLETADKQSVHIVENEIQASIDQIFSRLERLEILSSKEPPNKRQNARLRV
DQLKYDVQHLQTALRNFQHRRHAREQQERQREELLSRTFTTNDSDTTIPMDESLQFNSSLQKVHNGMDDL
ILDGHNILDGLRTQRLTLKGTQKKILDIANMLGLSNTVMRLIEKRAFQDKYFMIGGMLLTCVVMFLVVQY
LT

Human GOSR2 protein sequence - var3 (public gi: 12711467) (SEQ ID NO: 246)
MDPLFQQTHKQVHEIQSCMGRLETADKQSVHIVENEIQASIDQIFSRLERLEILSSKEPPNKRQNARLRV
DQLKYDVQHLQTALRNFQHRRHAREQQERQREELLSRTFTTNDSDTTIPMDESLQFNSSLQKVHNGMDDL
ILDGHNILDGLRTQRLTLKGTQKKILDIANMLGLSNTVMRLIEKRAFQDKYFMIGTQGSCQTAHFGGRSA
GSS

Human GOSR2 protein sequence - var4 (public gi: 21961349) (SEQ ID NO: 247)
MDPLFQQTHKQVHEIQSCMGRLETADKQSVHIVENEIQASIDQIFSRLERLEILSSKEPPNKRQNARLRV
DQLKYDVQHLQTALRNFQHRRHAREQQERQREELLSRTFTTNDSDTTIPMDESLQFNSSLQKVHNGMDDL
ILDGHNILDGLRTQRLTLKVGSLLGDREKASCFSLIQQFSNCVYILITCPQIVIF

Human GOSR2 protein sequence - var5 (public gi: 2316088) (SEQ ID NO: 248)
MDPLFQQTHKQVHEIQSCMGRLETADKQSVHIVENEIQASIDQIFSRLERLEILSSKEPPNKRQNARLRV
DQLKYDVQHLQTALRNFQHRRHAREQQERQREELLCRTFTTNGSDTTIPMDESLQFNSSLQKVHNGMDDL
ILDGHNILDGLRTQRLTLKGTQKKIPDIANMLGLSNTVMRLIEKRAFQDKYFMIGGMLLTCVVMFLVVQY
LT

Human GOSR2 pray sequence - var1 (SEQ ID NO: 76)
AGCGCCGCCATGGNAGTACCCATNCGACGTACCAGATTACGCTCATATGGCCATGGAGGCCAGTGAATTC
CACCCAAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCCGGAAACCGGAAGGGGGGCTGTGAGGACGT
GTTCCGAGGAAGCCAGACCCGGAGCCGTGGCCTGCCGGGCCGGCGACATGGATCCCCTGTTCCAGCAAAC
GCACAAGCAGGTCCACGAGATCCAGTCTTGCATGGGACGCCTGGAGACGGCAGACAAGCAGTCTGTGCAC
ATAGTAGAAAACGAAATCCAAGCAAGCATAGACCAGATATTCAGCCGTCTAGGACGTCTGGAGATTTTGT
CCAGCAAGGAGCCCCCTAACAAAAGGCAAAATGCCAAACTTCGGGTTGACCAGTTAAAGTATGATGTCCA
GCACCTGCAGACTGCGCTCAGAAACTTCCAGCATCGGCGCNATGCAAGGGAGCAGCGGGAGAGACAGCGA
GAAGANCTTNTGTCTCNAACCTTAACCNNTACCAANTTTGACNCCCCCTTNCCATTGACCAAATANTNGN
NGTTAACNTNCCTCCCNCNAAAAAGTTACAAACGGCTTGNNNAACNTANTTTAAAAGGGNCCNATTTTTT
TNAATNGCNTTGGGNNCCCAAAACCTTCCTTTNGNGGGGGGGNCCNTTTGGGGGGAAAAAAAAAANGCCC
TTTTTTTTANCCCCNNNNCAANNTTNAAANACNGNNNNTTNTTTTNAANCNGNNCCCCAAAGAGGGGAN
TTTTNNNAAANAAAAACNCCCCCCTNTGGGGGGGCCTTNTTTTGGGGNGGANNTTTTTGNNCCANNAAAA
ACCCCNTTTTNTNNGGNNGGAAAAAAAAAAAAAAGNNNNTNTTTNTA

Human HERPUD1 mRNA sequence - var1 (public gi: 16507801) (SEQ ID NO: 77)
AGAGACGTGAACGGTCGTTGCAGAGATTGCGGGCGGCTGAGACGCCGCCTGCCTGGCACCTAGGAGCGCA
GCGGAGCCCCGACACCGCCGCCGCCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTG
AAGAGCCCCAACCAGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCA
AGGCCCACCTGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAA
GCTGTTGTTGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGGAAAAACGGCATGTTTTGCATCTGGTG
TGCAATGTGAAGAGTCCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTG
CTGGTTCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAGGCAAAGGGAAGTTCTTCG
GAACCTTTCTTCCCCTGGATGGGAAAACATCTCAAGGCATCACGTTGGGTGGTTTCCATTTAGACCGAGG
CCGGTTCAGAACTTCCCAAATGATGGTCCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATAACTTAC
AGGAAGGCACTGATCCTGAAACTGAAGACCCCAACCACCTCCCTCCAGACAGGGATGTACTAGATGGCGA
GCAGACCAGCCCCTCCTTTATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTTCTTCCA
GAAGGCCCCCCAGCCATCGCAAACTGATGGTGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGA
CTGGATCACCTGACTCCAGCTAGATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGT
GGATGATGATATGCTTTTGTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAA
AAATGCCCAAGGCTTCTCATGTCTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCA
CTGTACGTAGAAGGCCTTAGGTGTTGCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTGTCTGCAT
GTGTGTTTGTACATAGAAGTCATAGATGCAGAAGTGGTTCTGCTGGTACGATTTGATTCCTGTTGGAATG
TTTAAATTACACTAAGTGTACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTT
CTAGGAAAGACTTATGTATAATTGCTTTTTAAAATGCAGTGCTTTACTTTAAACTAAGGGGAACTTTGCG
GAGGTGAAAACCTTTGCTGGGTTTTCTGTTCAATAAAGTTTTACTATGAATGACCCTGAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human HERPUD1 mRNA sequence - var2 (public gi: 10441910) (SEQ ID NO: 78)
GCTGTGTGGCCCAGGCTTTTCTCAAACTCCTGAGGGCAAGCGATCCTCCCACCTCAGCCTCCTGAGTAGC
TGGGACTACAGGCATGTGCCACTAGACCTGGCTCTAAAGACATATATGACACACGAAACCATTTATTTTT
CATTTCACAATGTTTATTCACATATATGGTATTAGTATTCTAATGTAGTGATGCACTCTAAATTTGCATT
ATATTTCCTAGAACATCTGAACAGAGCATAGGAAATTCCCTATTTTGCCATTATCAGTTCTAACAAAAAT
CTTAAAAGCACTTTATCATTTCATTTCCCTGCACTGTAATTTTTTTAAATGATCAAAAACAGTATCATAC
CAAGGCTTACTTATATTGGAATACTATTTTAGAAAGTTGTGGGCTGGGTTGTATTTATAAATCTTGTTGG
TCAGATGTCTGCAATGAGTAAATTTAGCACCATTATCAGGAAGCTTTCTCACCAATGACAACTTCATTGG
AAGATTTTAATGAAAGTGTAGCATACTCTAGGGAAAAAATATGAATATTTTAGCATCTATGTATTGAAAA
TTATGTTGAATAAATGTCAGACTATTTTTTACATAACGTTGCTTCTGTTTAATTTTGTCACGTTCAGAGG
TGGGGGGTAGGAGATGTAAGCCCTTGACAGCAAAATAATTCCTTTTGCTTGATTTCAGACAGTTGCATCA
GCTCCTTTGTTCTGTGTTCATGTTACACTTATTTAGGTGGCTGAATCCACAGAGGAGCCTGCTGGTTCTA
ATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAGGCAAAGGGAAGTTCTTCGGAACCTTTC
TTCCCCTGGATGGGAAAACATCTCAAGGCCTGAAGCTGCCCAGCAGGCATTCCAAGGCCTGGGTCCTGGT
TTCTCCGGTTACACACCCTATGGGTGGCTTCAGCTTTCCTGGTTCCAGCAGATATATGCACGACAGTACT
ACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCCACCACCAAGTGCACAAGAGATACC
TGTGGTCTCTGCACCTGCTCCAGCCCCTATTCACAACCAGTTTCCAGCTGAAAACCAGCCTGCCAATCAG
AATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGCACAAGGTGGCC
CTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTCAGCAGCTACATTTTC
TGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCCTCATGGTCATGGGGGCCACCGTT
GTTATGTACCTGCATCACGTTGGGTGGTTTCCATTTAGACCGAGGCCGGTTCAGAACTTCCCAAATGATG
GTCCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATAACTTACAGGAAGGCACTGATCCTGAAACTGA
AGACCCCAACCACCTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCCTCCTTTATGAGC
ACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTTCTTCCAGAAGGCCCCCCAGCCATCGCAAACT
GATGGTGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGACTCCAGCTAGAT
TGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGTGGATGATGATATGCTTTTGTGAGCA
AGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAAATGCCCAAGGCTTCTCATGTCTT
TATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCACTGTACGTAGAAGGCCTTAGGTGTT
GCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTGTCTGCATGTGTGTTTGTACATAGAAGTCATAG
ATGCAGAAGTGGTTCTGCTGGTACGATTTGATTCCTGTTGGAATGTTTAAATTACACTAAGTGTACTACT
TTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTATGTATAATTGC
TTTTTAAAATGCAGTGCTTTACTTTAAACTAAGGGGAACTTTGCGGAGGTGAAAACCTTTGCTGGGTTTT
CTGTTCAATAAAGTTTTACTATGAATGACAAAAAAAAAAAAAAAAAA

Human HERPUD1 mRNA sequence - var3 (public gi: 3005722) (SEQ ID NO: 79)
GGCCACCTCAAGGCCCACCTGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTT
ATTCTGGGAAGCTGTTGTTGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGGAAAAACGGCATGTTTT
GCATCTGGTGTGCAATGTGAAGAGTCCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACA
GAGGAGCCTGCTGGTTCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAGGCAAAGGG
AAGTTCTTCGGAACCTTTCTTCCCCTGGATGGGAAAACATCTCAAGGCCTGAAGCTGCCCAGCAGGCATT
CCAAGGCCTGGGTCCTGGTTTCTCCGGTTACACACCCTATGGGTGGCTTCAGCTTTCCTGGTTCCAGCAG
ATATATGCACGACAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCCACCAC
CAAGTGCACAAGAGATACCTGTGGTCTCTGCACCTGCTCCAGCCCCTATTCACAACCAGTTTCCAGCTGA
AAACCAGCCTGCCAATCAGAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGG
ATGAATGCACAAGGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCT
ATTCAGCAGCTACATTTTCTGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCCTCAT
GGTCATGGGGGCCACCGTTGTTATGTACCTGCATCACGTTGGGTGGTTTCCATTTAGACCGAGGCCGGTT
CAGAACTTCCCAAATGATGGTCCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATAACTTACAGGAAG
GCACTGATCCTGAAACTGAAGACCCCAACCACCTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGAC
CAGCCCCTCCTTTATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTTCTTCCAGAAGGC
CCCCCAGCCATCGCAAACTGATGGTGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGAT
CACCTGACTCCAGCTAGATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGTGGATGA
TGATATGCTTTTGTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAAATGC
CCAAGGCTTCTCATGTCTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCACTGTAC
GTAGAAGGCCTTAGGTGTTGCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTGTCTGCATGTGTGT
TTGTACATAGAAGTCATAGATGCAGAAGTGGTTCTGCTGGTACGATTTGATTCCTGTTGGAATGTTTAAA
TTACACTAAGTGTACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGA
AAGACTTATGTATAATTGCTTTTTAAAATGCAGTGCTTTACTTTAAACTAAGGGGAACTTTGCGGAGGTG
AAAACCTTTGCTGGGTTTTCTGTTCAATAAAGTTTTACTATGAATGACCCTGAAAAAAAAAAAAAAAAAA
AAAA

Human HERPUD1 mRNA sequence - var4 (public gi: 21619176) (SEQ ID NO: 80)
CCACGCGTCCGGGTCGTTGCAGAGATTGCGGGCGGCTGAGACGCCGCCTGCCTGGCACCTAGGAGCGCAG
CGGAGCCCCGACACCGCCGCCGCCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTGA
AGAGCCCCAACCAGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCAA
GGCCCACCTGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAAG
CTGTTGTTGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGCAGGAAAAACGGCATGTTTTGCATCTGG
TGTGCAATGTGAAGAGTCCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCC
TGCTGGTTCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAGGCAAAGGGAAGTTCTT
CGGAACCTTTCTTCCCCTGGATGGGAAAACATCTCAAGGCCTGAAGCTGCCCAGCAGGCATTCCAAGGCC
TGGGTCCTGGTTTCTCCGGTTACACACCCTATGGGTGGCTTCAGCTTTCCTGGTTCCAGCAGATATATGC
ACGACAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCCACCACCAAGTGCA
CAAGAGATACCTGTGGTCTCTGCACCTGCTCCAGCCCCTATTCACAACCAGTTTCCAGCTGAAAACCAGC
CTGCCAATCAGAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGC
ACAAGGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTCAGCA
GCTACATTTTCTGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCCTCATGGTCATGG
GGGCCACCGTTGTTATGTACCTGCATCACGTTGGGTGGTTTCCATTTAGACCGAGGCCGGTTCAGAACTT
CCCAAATGATGGTCCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATAACTTACAGGAAGGCACTGAT
CCTGAAACTGAAGACCCCAACCACCTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCCT
CCTTTATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTTCTTCCAGAAGGCCCCCCAGC
CATCGCAAACTGATGGTGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGAC
TCCAGCTAGATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGTGGATGATGATATGC
TTTTGTGAGCAAGCAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAAATGCCCAAGGCTT
CTCATGTCTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCACTGTACGTAGAAGGC
CTTAGGTGTTGCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTGTCTGCATGTGTGTTTGTACATA
GAAGTCATAGATGCAGAAGTGGTTCTGCTGGTACGATTTGATTCCTGTTGGAATGTTTAAATTACACTAA
GTGTACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTAT
GTATAATTGCTTTTTAAAATGCAGTGCTTTACTTTAAACTAAGGGGAACTTTGCGGAGGTGAAAACCTTT
GCTGGGTTTTCTGTTCAATAAAGTTTTACTATGAATGACCCTGAAAAAAAAAAAAAAA

Human HERPUD1 mRNA sequence - var5 (public gi: 14249882) (SEQ ID NO: 81)
AACGGTCGTTGCAGAGATTGCGGGCGGCTGAGACGCCGCCTGCCTGGCACCTAGGAGCGCAGCGGAGCCC
CGACACCGCCGCCGCCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTGAAGAGCCCC
AACCAGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCAAGGCCCACC
TGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAAGCTGTTGTT
GGATCACCAATGTCTCAGGGACTTGCTTCCAAAGCAGGAAAAACGGCATGTTTTGCATCTGGTGTGCAAT
GTGAAGAGTCCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTGCTGGTT
CTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAGGCAAAGGGAAGTTCTTCGGAACCT
TTCTTCCCCTGGATGGGAAAACATCTCAAGGCCTGAAGCTGCCCAGCAGGCATTCCAAGGCCTGGGTCCT
GGTTTCTCCGGTTACACACCCTATGGGTGGCTTCAGCTTTCCTGGTTCCAGCAGATATATGCACGACAGT
ACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCCACCACCAAGTGCACAAGAGAT
ACCTGTGGTCTCTGCACCTGCTCCAGCCCCTATTCACAACCAGTTTCCAGCTGAAAACCAGCCTGCCAAT
CAGAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGCACAAGGTG
GCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTCAGCAGCTACATT
TTCTGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCCTCATGGTCATGGGGGCCACC
GTTGTTATGTACCTGCATCACGTTGGGTGGTTTCCATTTAGACCGAGGCCGGTTCAGAACTTCCCAAATG
ATGGTCCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATAACTTACAGGAAGGCACTGATCCTGAAAC
TGAAGACCCCAACCACCTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCCTCCTTTATG
AGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTTCTTCCAGAAGGCCCCCCAGCCATCGCAA
ACTGATGGTGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGACTCCAGCTA
GATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGTGGATGATGATATGCTTTTGTGA
GCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAAATGCCCAAGGCTTCTCATGT
CTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCACTGTACGTAGAAGGCCTTAGGT
GTTGCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTGTCTGCATGTGTGTTTGTACATAGAAGTCA
TAGATGCAGAAGTGGTTCTGCTGGTACGATTTGATTCCTGTTGGAATGTTTAAATTACACTAAGTGTACT
ACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTATGTATAAT
TGCTTTTTAAAATGCAGTGCTTTACTTTAAACTAAGGGGAACTTTGCGGAGGTGAAAACCTTTGCTGGGT
TTTCTGTTCAATAAAGTTTTACTATGAAAAAAAAAAAAAAAAAA

Human HERPUD1 mRNA sequence - var6 (public gi: 12652674) (SEQ ID NO: 82)
GAACTGTCGTTGCAGAGATTGCGGGCGGCTGAGACGCCGCCTGCCTGGCACCTAGGAGCGCAGCGGAGCC
CCGACACCGCCGCCGCCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTGAAGAGCCC
CAACCAGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCAAGGCCCAC

CTGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAAGCTGTTGT
TGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGCAGGAAAAACGGCATGTTTTGCATCTGGTGTGCAA
TGTGAAGAGTCCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTGCTGGT
TCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAGGCAAAGGGAAGTTCTTCGGAACC
TTTCTTCCCCTGGATGGGAAAACATCTCAAGGCCTGAAGCTGCCCAGCAGGCATTCCAAGGCCTGGGTCC
TGGTTTCTCCGGTTACACACCCTATGGGTGGCTTCAGCTTTCCTGGTTCCAGCAGATATATGCACGACAG
TACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCCACCACCAAGTGCACAAGAGA
TACCTGTGGTCTCTGCACCTGCTCCAGCCCCTATTCACAACCAGTTTCCAGCTGAAAACCAGCCTGCCAA
TCAGAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGCACAAGGT
GGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTCAGCAGCTACAT
TTTCTGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCCTCATGGTCATGGGGGCCAC
CGTTGTTATGTACCTGCATCACGTTGGGTGGTTTCCATTTAGACCGAGGCCGGTTCAGAACTTCCCAAAT
GATGGTCCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATAACTTACAGGAAGGCACTGATCCTGAAA
CTGAAGACCCCAACCACCTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCCTCCTTTAT
GAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTTCTTCCAGAAGGCCCCCCAGCCATCGCA
AACTGATGGTGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGACTCCAGCT
AGATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGTGGATGATGATATGCTTTTGTG
AGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAAATGCCCAAGGCTTCTCATG
TCTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCACTGTACGTAGAAGGCCTTAGG
TGTTGCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTGTCTGCATGTGTGTTTGTACATAGAAGTC
ATAGATGCAGAAGTGGTTCTGCTGGTACGATTTGATTCCTGTTGGAATGTTTAAATTACACTAAGTGTAC
TACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTATGTATAA
TTGCTTTTTAAAATGCAGTGCTTTACTTTAAACTAAGGGGAACTTTGCGGAGGTGAAAACCTTTGCTGGG
TTTTCTGTTCAATAAAGTTTTACTATGAATGAAAAAAAAAAAAAAAAAAAA

Human HERPUD1 mRNA sequence - var7 (public gi: 9711684) (SEQ ID NO: 83)
AGAGACGTGAACTGTCGTTGCAGAGATTGCGGGCGGCTGAGACGCCGCCTGCCTGGCACCTAGGAGCGCA
GCGGAGCCCCGACACCGCCGCCGCCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTG
AAGAGCCCCAACCAGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCA
AGGCCCACCTGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAA
GCTGTTGTTGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGCAGGAAAAACGGCATGTTTTGCATCTG
GTGTGCAATGTGAAGAGTCCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGC
CTGCTGGTTCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAGGCAAAGGGAAGTTCT
TCGGAACCTTTCTTCCCCTGGATGGGAAAACATCTCAAGGCCTGAAGCTGCCCAGCAGGCATTCCAAGGC
CTGGGTCCTGGTTTCTCCGGTTACACACCCTATGGGTGGCTTCAGCTTTCCTGGTTCCAGCAGATATATG
CACGACAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCCACCACCAAGTGC
ACAAGAGATACCTGTGGTCTCTGCACCTGCTCCAGCCCCTATTCACAACCAGTTTCCAGCTGAAAACCAG
CCTGCCAATCAGAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATG
CACAAGGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTCAGC
AGCTACATTTTCTGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCCTCATGGTCATG
GGGGCCACCGTTGTTATGTACCTGCATCACGTTGGGTGGTTTCCATTTAGACCGAGGCCGGTTCAGAACT
TCCCAAATGATGGTCCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATAACTTACAGGAAGGCACTGA
TCCTGAAACTGAAGACCCCAACCACCTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCC
TCCTTTATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTTCTTCCAGAAGGCCCCCCAG
CCATCGCAAACTGATGGTGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGA
CTCCAGCTAGATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGTGGATGATGATATG
CTTTTGTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAAATGCCCAAGGC
TTCTCATGTCTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCACTGTACGTAGAAG
GCCTTAGGTGTTGCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTGTCTGCATGTGTGTTTGTACA
TAGAAGTCATAGATGCAGAAGTGGTTCTGCTGGTACGATTTGATTCCTGTTGGAATGTTTAAATTACACT
AAGTGTACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTT
ATGTATAATTGCTTTTTAAAATGCAGTGCTTTACTTTAAACTAAGGGGAACTTTGCGGAGGTGAAAACCT
TTGCTGGGTTTTCTGTTCAATAAAGTTTTACTATGAATGACCCTG

Human HERPUD1 mRNA sequence - var8 (public gi: 3005718) (SEQ ID NO: 84)
GACGTGAACGGTCGTTGCAGAGATTGCGGGCGGCTGAGACGCCGCCTGCCTGGCACCTAGGAGCGCAGCG
GAGCCCCGACACCGCCGCCGCCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTGAAG
AGCCCCAACCAGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCAAGG
CCCACCTGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAAGCT
GTTGTTGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGCAGGAAAAACGGCATGTTTTGCATCTGGTG
TGCAATGTGAAGAGTCCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTG
CTGGTTCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAGGCAAAGGGAAGTTCTTCG

GAACCTTTCTTCCCCTGGATGGGAAAACATCTCAAGGCCTGAAGCTGCCCAGCAGGCATTCCAAGGCCTG
GGTCCTGGTTTCTCCGGTTACACACCCTATGGGTGGCTTCAGCTTTCCTGGTTCCAGCAGATATATGCAC
GACAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCCACCACCAAGTGCACA
AGAGATACCTGTGGTCTCTGCACCTGCTCCAGCCCCTATTCACAACCAGTTTCCAGCTGAAAACCAGCCT
GCCAATCAGAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGCAC
AAGGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTCAGCAGC
TACATTTTCTGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCCTCATGGTCATGGGG
GCCACCGTTGTTATGTACCTGCATCACGTTGGGTGGTTTCCATTTAGACCGAGGCCGGTTCAGAACTTCC
CAAATGATGGTCCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATAACTTACAGGAAGGCACTGATCC
TGAAACTGAAGACCCCAACCACCTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCCTCC
TTTATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTTCTTCCAGAAGGCCCCCCAGCCA
TCGCAAACTGATGGTGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGACTC
CAGCTAGATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGTGGATGATGATATGCTT
TTGTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAAATGCCCAAGGCTTC
TCATGTCTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCACTGTACGTAGAAGGCC
TTAGGTGTTGCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTGTCTGCATGTGTGTTTGTACATAG
AAGTCATAGATGCAGAAGTGGTTCTGCTGGTACGATTTGATTCCTGTTGGAATGTTTAAATTACACTAAG
TGTACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTATG
TATAATTGCTTTTTAAAATGCAGTGCTTTACTTTAAACTAAGGGGAACTTTGCGGAGGTGAAAACCTTTG
CTGGGTTTTCTGTTCAATAAAGTTTTACTATGAATGACCCTGAAAAAAAAAAAAAAAAAAAAA

Human HERPUD1 mRNA sequence - var9 (public gi: 285960) (SEQ ID NO: 85)
CGTGAACGGTCGTTGCAGAGATTGCGGGCGGCTGAGACGCCGCCTGCCTGGCACCTAGGAGCGCAGCGGA
GCCCCGACACCGCCGCCGCCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTGAAGAG
CCCCAACCAGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCAAGGCC
CACCTGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAAGCTGT
TGTTGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGCAGGAAAAACGGCATGTTTTGCATCTGGTGTG
CAATGTGAAGAGTCCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTGCT
GGTTCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAGGCAAAGGGAAGTTCTTCGGA
ACCTTTCTTCCCCTGGATGGGAAAACATCTCAAGGCCTGAAGCTGCCCAGCAGGCATTCCAAGGCCTGGG
TCCTGGTTTCTCCGGTTACACACCCTATGGGTGGCTTCAGCTTTCCTGGTTCCAGCAGATATATGCACGA
CAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCCACCACCAAGTGCACAAG
AGATACCTGTGGTCTCTGCACCTGCTCCAGCCCCTATTCACAACCAGTTTCCAGCTGAAAACCAGCCTGC
CAATCAGAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGCACAA
GGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTCAGCAGCTA
CATTTTCTGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCCTCATGGTCATGGGGGC
CACCGTTGTTATGTACCTGCATCACGTTGGGTGGTTTCCATTTAGACCGAGGCCGGTTCAGAACTTCCCA
AATGATGGTCCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATAACTTACAGGAAGGCACTGATCCTG
AAACTGAAGACCCCAACCACCTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCCTCCTT
TATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTTCTTCCAGAAGGCCCCCCAGCCATC
GCAAACTGATGGTGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGACTCCA
GCTAGATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGTGGATGATGATATGCTTTT
GTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAAATGCCCAAGGCTTCTC
ATGTGTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCACTGTACGTAGAAGGCCTT
AGGTGTTGCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTGTCTGCATGTGTGTTTGTACATAGAA
GTCATAGATGCAGAAGTGGTTCTGCTGGTAAGATTTGATTCCTGTTGGAATGTTTAAATTACACTAAGTG
TACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTATGTA
TAATTGCTTTTTAAAATGCAGTGCTTTACTTTAAACTAAGGGGAACTTTGCGGAGGTGAAAACCTTTGCT
GGGTTTTCTGTTCAATAAAGTTTTACTATGAATGACCCTG

Human HERPUD1 mRNA sequence - var10 (public gi: 7661869) (SEQ ID NO: 86)
GACGTGAACGGTCGTTGCAGAGATTGCGGGCGGCTGAGACGCCGCCTGCCTGGCACCTAGGAGCGCAGCG
GAGCCCCGACACCGCCGCCGCCGCCATGGAGTCCGAGACCGAACCCGAGCCCGTCACGCTCCTGGTGAAG
AGCCCCAACCAGCGCCACCGCGACTTGGAGCTGAGTGGCGACCGCGGCTGGAGTGTGGGCCACCTCAAGG
CCCACCTGAGCCGCGTCTACCCCGAGCGTCCGCGTCCAGAGGACCAGAGGTTAATTTATTCTGGGAAGCT
GTTGTTGGATCACCAATGTCTCAGGGACTTGCTTCCAAAGCAGGAAAAACGGCATGTTTTGCATCTGGTG
TGCAATGTGAAGAGTCCTTCAAAAATGCCAGAAATCAACGCCAAGGTGGCTGAATCCACAGAGGAGCCTG
CTGGTTCTAATCGGGGACAGTATCCTGAGGATTCCTCAAGTGATGGTTTAAGGCAAAGGGAAGTTCTTCG
GAACCTTTCTTCCCCTGGATGGGAAAACATCTCAAGGCCTGAAGCTGCCCAGCAGGCATTCCAAGGCCTG
GGTCCTGGTTTCTCCGGTTACACACCCTATGGGTGGCTTCAGCTTTCCTGGTTCCAGCAGATATATGCAC
GACAGTACTACATGCAATATTTAGCAGCCACTGCTGCATCAGGGGCTTTTGTTCCACCACCAAGTGCACA
AGAGATACCTGTGGTCTCTGCACCTGCTCCAGCCCCTATTCACAACCAGTTTCCAGCTGAAAACCAGCCT

GCCAATCAGAATGCTGCTCCTCAAGTGGTTGTTAATCCTGGAGCCAATCAAAATTTGCGGATGAATGCAC
AAGGTGGCCCTATTGTGGAAGAAGATGATGAAATAAATCGAGATTGGTTGGATTGGACCTATTCAGCAGC
TACATTTTCTGTTTTTCTCAGTATCCTCTACTTCTACTCCTCCCTGAGCAGATTCCTCATGGTCATGGGG
GCCACCGTTGTTATGTACCTGCATCACGTTGGGTGGTTTCCATTTAGACCGAGGCCGGTTCAGAACTTCC
CAAATGATGGTCCTCCTCCTGACGTTGTAAATCAGGACCCCAACAATAACTTACAGGAAGGCACTGATCC
TGAAACTGAAGACCCCAACCACCTCCCTCCAGACAGGGATGTACTAGATGGCGAGCAGACCAGCCCCTCC
TTTATGAGCACAGCATGGCTTGTCTTCAAGACTTTCTTTGCCTCTCTTCTTCCAGAAGGCCCCCCAGCCA
TCGCAAACTGATGGTGTTTGTGCTGTAGCTGTTGGAGGCTTTGACAGGAATGGACTGGATCACCTGACTC
CAGCTAGATTGCCTCTCCTGGACATGGCAATGATGAGTTTTTAAAAAACAGTGTGGATGATGATATGCTT
TTGTGAGCAAGCAAAAGCAGAAACGTGAAGCCGTGATACAAATTGGTGAACAAAAAATGCCCAAGGCTTC
TCATGTCTTTATTCTGAAGAGCTTTAATATATACTCTATGTAGTTTAATAAGCACTGTACGTAGAAGGCC
TTAGGTGTTGCATGTCTATGCTTGAGGAACTTTTCCAAATGTGTGTGTCTGCATGTGTGTTTGTACATAG
AAGTCATAGATGCAGAAGTGGTTCTGCTGGTACGATTTGATTCCTGTTGGAATGTTTAAATTACACTAAG
TGTACTACTTTATATAATCAATGAAATTGCTAGACATGTTTTAGCAGGACTTTTCTAGGAAAGACTTATG
TATAATTGCTTTTTAAAATGCAGTGCTTTACTTTAAACTAAGGGGAACTTTGCGGAGGTGAAAACCTTTG
CTGGGTTTTCTGTTCAATAAAGTTTACTATGAATGACCCTGAAAAAAAAAAAAAAAAAAAAA

Human HERPUD1 Protein sequence - var1 (public gi: 16507802) (SEQ ID NO: 249)
MESETEPEPVTLLVKSPNQRHRDLELSGDRGWSVGHLKAHLSRVYPERPRPEDQRLIYSGKLLLDHQCLR
DLLPKEKRHVLHLVCNVKSPSKMPEINAKVAESTEEPAGSNRGQYPEDSSSDGLRQREVLRNLSSPGWEN
ISRHHVGWFPFRPRPVQNFPNDGPPPDVVNQDPNNNLQEGTDPETEDPNHLPPDRDVLDGEQTSPSFMST
AWLVFKTFFASLLPEGPPAIAN

Human HERPUD1 Protein sequence - var2 (public gi: 10441911) (SEQ ID NO: 250)
MQYLAATAASGAFVPPPSAQEIPVVSAPAPAPIHNQFPAENQPANQNAAPQVVVNPGANQNLRMNAQGGP
IVEEDDEINRDWLDWTYSAATFSVFLSILYFYSSLSRFLMVMGATVVMYLHHVGWFPFRPRPVQNFPNDG
PPPDVVNQDPNNNLQEGTDPETEDPNHLPPDRDVLDGEQTSPSFMSTAWLVFKTFFASLLPEGPPAIAN

Human HERPUD1 Protein sequence - var3 (public gi: 3005723) (SEQ ID NO: 251)
GHLKAHLSRVYPERPRPEDQRLIYSGKLLLDHQCLRDLLPKEKRHVLHLVCNVKSPSKMPEINAKVAEST
EEPAGSNRGQYPEDSSSDGLRQREVLRNLSSPGWENISRPEAAQQAFQGLGPGFSGYTPYGWLQLSWFQQ
IYARQYYMQYLAATAASGAFVPPPSAQEIPVVSAPAPAPIHNQFPAENQPANQNAAPQVVVNPGANQNLR
MNAQGGPIVEEDDEINRDWLDWTYSAATFSVFLSILYFYSSLSRFLMVMGATVVMYLHHVGWFPFRPRPV
QNFPNDGPPPDVVNQDPNNNLQEGTDPETEDPNHLPPDRDVLDGEQTSPSFMSTAWLVFKTFFASLLPEG
PPAIAN

Human HERPUD1 Protein sequence - var4 (public gi: 7661870) (SEQ ID NO: 252)
MESETEPEPVTLLVKSPNQRHRDLELSGDRGWSVGHLKAHLSRVYPERPRPEDQRLIYSGKLLLDHQCLR
DLLPKQEKRHVLHLVCNVKSPSKMPEINAKVAESTEEPAGSNRGQYPEDSSSDGLRQREVLRNLSSPGWE
NISRPEAAQQAFQGLGPGFSGYTPYGWLQLSWFQQIYARQYYMQYLAATAASGAFVPPPSAQEIPVVSAP
APAPIHNQFPAENQPANQNAAPQVVVNPGANQNLRMNAQGGPIVEEDDEINRDWLDWTYSAATFSVFLSI
LYFYSSLSRFLMVMGATVVMYLHHVGWFPFRPRPVQNFPNDGPPPDVVNQDPNNNLQEGTDPETEDPNHL
PPDRDVLDGEQTSPSFMSTAWLVFKTFFASLLPEGPPAIAN

Unigene Name: HLA-A Unigene ID: Hs.181244 Clone ID: GD_159

Human HLA-A mRNA sequence - var1 (public gi: 575248) (SEQ ID NO: 87)
ATGGCCGTCATGGCGCCCCGAACCCTCGTCCTGCTACTCTCGGGGGCTCTGGCCCTGACCCAGACCTGGG
CGGGCTCTCACTCCATGAGGTATTTCTTCACATCCGTGTCCCGGCCCGGCCGCGGGGAGCCCCGCTTCAT
CGCAGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACGCCGCGAGCCAGAGGATGGAG
CCGCGGGCGCCGTGGATAGAGCAGGAGGGTCCGGAGTATTGGGACGGGGAGACACGGAAAGTGAAGGCCC
ACTCACAGACTCACCGAGTGGACCTGGGGACCCTGCGCGGCTACTACAACCAGAGCGAGGCCGGTTCTCA
CACCGTCCAGAGGATGTATGGCTGCGACGTGGGGTCGGACTGGCGCTTCCTCCGCGGGTACCACCAGTAC
GCCTACGACGGCAAGGATTACATCGCCCTGAAAGAGGACCTGCGCTCTTGGACCGCGGCGGACATGGCAG
CTCAGACCACCAAGCACAAGTGGGAGGCGGCCCATGTGGCGGAGCAGTTGAGAGCCTACCTGGAGGGCGA
GTGCGTGGAGTGGCTCCGCAGATACCTGGAGAACGGGAAGGAGACGCTGCAGCGCACGGACGCCCCCAAA
ACGCATATGACTCACCACGCTGTCTCTGACCATGAAGCCACCCTGAGGTGCTGGGCCCTGAGCTTCTACC

CTGCGGAGATCACACTGACCTGGCAGCGGGATGGGGAGGACCAGACCCAGGACACGGAGCTCGTGGAGAC
CAGGCCTGCAGGGGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAGA
TACACCTGCCATGTGCAGCATGAGGGTTTGCCCAAGCCCCTCACCCTGAGATGGGAGCCGTCTTCCCAGC
CCACCATCCCCATCGTGGGCATCATTGCTGGCCTGGTTCTCTTTGGAGCTGTGATCACTGGAGCTGTGGT
CGCTGCTGTGATGTGGAGGAGGAAGAGCTCAGATAGAAAAGGAGGGAGCTACTCTCAGGCTGCAAGCAGT
GACAGTGCCCAGGGCTCTGATGTGTCTCTCACAGCTTGTAAAGTGTGA

Human HLA-A mRNA sequence - var2 (public gi: 187857) (SEQ ID NO: 88)
ATGGCCGTCATGGCGCCCCGAACCCTCGTCCTGCTACTCTCGGGGGCCCTGGCCCTGACCCAGACCTGGG
CGGGCTCCCACTCCATGAGGTATTTCTACACTTCCGTGTCCCGGCCCGGCCGCGGGGAGCCCCGCTTCAT
CGCCGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACGCCGCGAGCCAGAGGATGGAG
CCGCGGGCGCCGTGGATAGAGCAGGAGGGGCCGGAGTATTGGGACCGGAACACACGGAATGTGAAGGCCC
AGTCACAGACTGACCGAGTGGACCTGGGGACCCTGCGCGGCTACTACAACCAGAGCGAGGCCGGTTCTCA
CACCATCCAGATGATGTATGGCTGCGACGTGGGGTCGGACGGGCGCTTCCTCCGCGGGTACCGGCAGGAC
GCCTACGACGGCAAGGATTACATCGCCCTGAAAGAGGACCTGCGCTCTTGGACCGCGGCGGACATGGCAG
CTCAGACCACCAAGCACAAGTGGGAGGCGGCCCATGTGGCGGAGCAGTGGAGAGCCTACCTGGAGGGCAC
GTGCGTGGAGTGGCTCCGCAGATACCTGGAGAACGGGAAGGAGACGCTGCAGCGCACGGACGCCCCCAAA
ACGCATATGACTCACCACGCTGTCTCTGACCATGAAGCCACCCTGAGGTGCTGGGCCCTGAGCTTCTACC
CTGCGGAGATCACACTGACCTGGCAGCGGGATGGGGAGGACCAGACCCAGGACACGGAGCTCGTGGAGAC
CAGGCCTGCAGGGGATGGAACCTTCCAGAAGTGGGTGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAGA
TACACCTGCCATGTGCAGCATGAGGGTTTGCCCAAGCCCCTCACCCTGAGATGGGAGCCGTCTTCCCAGC
CCACCATCCCCATCGTGGGCATCATTGCTGGCCTGGTTCTCTTTGGAGCTGTGATCACTGGAGCTGTGGT
CGCTGCTGTGATGTGGAGGAGGAAGAGCTCAGATAGAAAAGGAGGGAGCTACTCTCAGGCTGCAAGCAGT
GACAGTGCCCAGGGCTCTGATGTGTCTCTCACAGCTTGTAAAGTGTGA

Human HLA-A protein sequence - var1 (public gi: 575249) (SEQ ID NO: 253)
MAVMAPRTLVLLLSGALALTQTWAGSHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRME
PRAPWIEQEGPEYWDGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQY
AYDGKDYIALKEDLRSWTAADMAAQTTKHKWEAAHVAEQLRAYLEGECVEWLRRYLENGKETLQRTDAPK
THMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGEDQTQDTELVETRPAGDGTFQKWAAVVVPSGQEQR
YTCHVQHEGLPKPLTLRWEPSSQPTIPIVGIIAGLVLFGAVITGAVVAAVMWRRKSSDRKGGSYSQAASS
DSAQGSDVSLTACKV

Unigene Name: HLA-B **Unigene ID**: Hs.77961 **Clone ID**: 3GD_1122

Human HLA-B mRNA sequence - var1 (public gi: 32188) (SEQ ID NO: 89)
ATGCGGGTCACGGCGCCCCGAACCGTCCTCCTGCTGCTCTCGGGAGCCCTGGCCCTGACCGAGACCTGGG
CCGGCTCCCACTCCATGAGGTATTTCTACACCGCCATGTCCCGGCCCGGCCGCGGGGAGCCCCGCTTCAT
CTCAGTGGGCTACGTGGACGACACGCAGTTCGTGAGGTTCGACAGCGACGCCGCGAGTCCGAGAGAGGAG
CCGCGGGCGCCGTGGATAGAGCAGGAGGGGCCGGAGTATTGGGACCGGGAGACACAGATCTCCAAGACCA
ACACACAGACTTACCGAGAGAGCCTGCGGAACCTGCGCGGCTACTACAACCAGAGCGAGGCCGGGTCTCA
CACCCTCCAGAGGATGTACGGCTGCGACGTGGGGCCGGACGGGCGCCTCCTCCGCGGGCATGACCAGTCC
GCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGAGCTCCTGGACCGCGGCGGACACGGCGG
CTCAGATCACCCAGCGCAAGTGGGAGGCGGCCCGTGAGGCGGAGCAGCTGAGAGCCTACCTGGAGGGCCT
GTGCGTGGAGTGGCTCCGCAGATACCTGGAGAACGGGAAGGAGACGCTGCAGCGCGCGGACCCCCCAAAG
ACACATGTGACCCACCACCCCATCTCTGACCATGAGGCCACCCTGAGGTGCTGGGCCCTGGGCTTCTACC
CTGCGGAGATCACACTGACCTGGCAGCGGGATGGCGAGGACCAAACTCAGGACACCGAGCTTGTGGAGAC
CAGACCAGCAGGAGATAGAACCTTCCAGAAGTGGGCAGCTGTGGTGGTGCCTTCTGGAGAAGAGCAGAGA
TACACATGCCATGTACAGCATGAGGGGCTGCCGAAGCCCCTCACCCTGAGATGGGAGCCATCTTCCCAGT
CCACCATCCCCATCGTGGGCATTGTTGCTGGCCTGGCTGTCCTAGCAGTTGTGGTCATCGGAGCTGTGGT
CGCTACTGTGATGTGTAGGAGGAAGAGCTCAGGTGGAAAAGGAGGGAGCTACTCTCAGGCTGCGTCCAGC
GACAGTGCCCAGGGCTCTGATGTGTCTCTCACAGCTTGA

Human HLA-B protein sequence - var1 (public gi: 32189) (SEQ ID NO: 254)
MRVTAPRTVLLLLSGALALTETWAGSHSMRYFYTAMSRPGRGEPRFISVGYVDDTQFVRFDSDAASPREE
PRAPWIEQEGPEYWDRETQISKTNTQTYRESLRNLRGYYNQSEAGSHTLQRMYGCDVGPDGRLLRGHDQS
AYDGKDYIALNEDLSSWTAADTAAQITQRKWEAAREAEQLRAYLEGLCVEWLRRYLENGKETLQRADPPK
THVTHHPISDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTELVETRPAGDRTFQKWAAVVVPSGEEQR
YTCHVQHEGLPKPLTLRWEPSSQSTIPIVGIVAGLAVLAVVVIGAVVATVMCRRKSSGGKGGSYSQAASS
DSAQGSDVSLTA

Unigene Name: MSTP028 **Unigene ID**: Hs.302746 **Clone ID**: GD_1119

Human MSTP028 mRNA sequence - var1 (public gi: 14042294) (SEQ ID NO: 90)
CCCCGCCTCCGCCCCCGGCTGGCGTGAGCTGGGTGTTTCCTGCCTCTCTCAGTCCGGGTTTGGAGACTCC
TGCGTCCTCCGACTTTTCGTGGAAGAGATGTCAGGAGAAAGTGTGGTGAGCTCAGCGGTGCCAGCGGCTG
CTACCCGCACCACTTCCTTCAAGGGCACGAGCCCCAGCTCCAAATACGTGAAGCTGAATGTGGGTGGAGC
CCTCTACTATACCACCATGCAGACGCTGACCAAGCAGGACACCATGCTGAAGGCCATGTTCAGCGGGCGC
ATGGAAGTGCTCACCGACAGTGAAGGCTGGATCCTCATTGACCGCTGTGGGAAGCACTTTGGTACGATAC
TCAACTACCTTCGAGACGGGGCGGTGCCTTTACCCGAGAGCCGCCGGGAGATCGAGGAGCTGCTAGCAGA
AGCCAAGTACTACCTAGTCCAAGGCCTGGTGGAAGAGTGCCAGGCGGCCCTACAAAACAAAGATACTTAT
GAGCCTTTCTGCAAGGTCCCTGTGATCACCTCATCCAAGGAAGAACAAAAACTTATAGCGACTTCAAATA
AGCCAGCCGTGAAGTTGCTCTACAACAGAAGTAACAACAAATACTCATATACCAGCAATTCTGACGACAA
TATGTTGAAAAACATTGAACTGTTTGATAAGCTGTCTCTGCGCTTTAACGGAAGGGTCCTGTTCATAAAG
GATGTTATTGGGGATGAAATCTGCTGCTGGTCCTTTTATGGTCAGGGCCGGAAGATTGCTGAAGTCTGTT
GTACCTCCATCGTCTATGCCACTGAGAAGAAACAGACCAAGGTGGAGTTTCCCGAAGCCCGGATTTATGA
GGAGACCCTGAACATTTTGCTGTATGAGGCCCAGGATGGCCGGGGACCTGACAATGCGCTCCTGGAGGCC
ACAGGCGGGGCGGCGGGGCGCTCCCACCACCTGGACGAGGACGAGGAGCGGGAGCGGATCGAGCGCGTGC
GGAGGATCCACATCAAGCGCCCTGATGACCGGGCCCACCTCCACCAGTGAGCAGGCAAGAGACCGAGCCG
CCCTCCTCTCACCGCCCCCACTCCCTGCCGTGCTACACCCAGATCCTGTGCAGGCTGCCGGGCCCCCTTCT
GCTTCCCTTGGAGCCTGGAGATACTTTTGTAACAAGCCAGATGATTATTTTGGTATTGCTTGACAAGGCA
AATTGATTGTCTTGACCCAGGCGTATGACCCCTGTCGTTGAACAAGCTGTGTCTAAGATCTCTACTTTTC
ATGAGAATCTGAGACTCTTTGGAGCCAGGCTTTCTCGGTTCTCAGAGGAAAAGTATGAATGAGTGTGAAG
TGTATGTGAGAACTTTTGTTTGCAATATTTATTTTTGTGGGTGTCGGCTTCCTATGTGGGCTTTTTGGGT
GACACTCCCTTAAGGGTTCAGTTTGACAATTCTGAGAGTTGTCCTGCAGTTGGAGGCCACCAGAGGTATC
TGAGCTCCCTGCTTCCTATTTCATAATCCTCCAGCCCCAGCAGGTCCACTCCTGGTTCCTGTGTGTTTGG
CCCGGGCACAATCCCCACTGCTTTGCTAGACGTGCTTTCTGCCATGTGGCTTTGGGCCTAGAGCTTGTTG
ATAATTGCAGCTTGTGGCAGGGGAAATATGGCTGAATGAGCGTCTAAATCGTTGAGACCAGTGCAACTTT
GGGTGCAAGGCTTTGTTTAGGGATCAAGCCTTTTGCCACCTTGGGCTGGTCTTTGGCCTGGTGCTCACTG
GGACCCCATATGTCTGCGTAGGAGCAGAACTTTCCATGGCAGTAAGTGTCCAGCTCTGTTTCTGGTTCTT
TCCCCAACTCCAGCCCCGTCCAGTTGTTCTCCTGATTGACCCGACTCCACTCCAGGAAGGCCATCTGACC

CTGTGACAGGCATAGCTCATAAACTACCCCTCCCTGGGATCCCGCTCCTCTTCAGCCTCCTTCCCCATGA
AGCTGGGCTAACTTTCTAAGTCATTTGCTTAGAAATTCAGTGTGGCCCATACCCTTTGTCCTCCCAGCC
TGGCATCCAGGCAGGGACACCCTCACACCACCAGCCCCAGGGAGCTTCCCTGCTATAAACACAGACCCCC
TTGTCTTTGCCTCTGATTTTTACACAGTGTAGAGTGGCCAGCAGTGAACAGGTTGAGGATGTGCGGGTAG
ATAGATAACTTTGGGTCTGGTTTGTGTCTGTGTTCATGTTCGTTTAAGGGATATGTGTGACTGTGGGTGG
GGACGTGTGCTTGTGGGGCACAGGTGGCGGCCCCTGCTGGAGCCCGGCTGGGCGCAGCGCCTATGTAGGA
CGGGTGTTCTCAGTGACCTACCTCCCAGGCTCCTCTGCACCTGCAAAGGAACAGGAGTGAGTCGTGACTG
ACAGGGGTGGTTGAGACTAGACTAGGTAGAGTAGTTACCAGGAGATGTGAATGTGCGTCAGGTGATGGAT
GGGTTTGTCAAGGGAATCGTTACCGTTTTATACCAAAGGTATTAACATGGGCAGCCTTTGACACATGTAT
TCCAAAAACGAGTTTATATTTTCAAACGGTTTTTACAGCTTAGACTTTGTACTTACTGCCCTGCCTGTGA
CAGTTGTATGCCTTCATTTGTATCCAACAGCAAAGTCTACAATAAAACTTTAAAACAATCATG

Human MSTP028 mRNA sequence - var2 (public gi: 13994352) (SEQ ID NO: 91)
GGAGACTCCTGCGTCCTCCGACTTTTCATGGAAGAGATGTCAGGAGAAAGTGTGGTGAGCTCAGCGGTGC
CAGCGGCTGCTACCCGCACCACTTCCTTCAAGGGCACGAGCCCCAGCTCCAAATACGTGAAGCTGAATGT
GGGTGGAGCCCTCTACTATACCACCATGCAGACGCTGACCAAGCAGGACACCATGCTGAAGGCCATGTTC
AGCGGGCGCATGGAAGTGCTCACCGACAGTGAAGGCTGGATCCTCATTGACCGCTGTGGGAAGCACTTTG
GTACGATACTCAACTACCTTCGAGACGGGGCGGTGCCTTTACCCGAGAGCCGCCGGGAGATCGAGGAGCT
GCTAGCAGAAGCCAAGTACTACCTAGTCCAAGGCCTGGTGGAAGAGTGCCAGGCGGCCCTACAAAACAAA
GATACTTATGAGCCTTTCTGCAAGGTCCCTGTGATCACCTCATCCAAGGAAGAACAAAAACTTATAGCGA
CTTCAAATAAGCCAGCCGTGAAGTTGCTCTACAACAGAAGTAACAACAAATACTCATATACCAGCAATTC
TGACGACAATATGTTGAAAAACATTGAACTGTTTGATAAGCTGTCTCTGCGCTTTAACGGAAGGGTCCTG
TTCATAAAGGATGTTATTGGGGATGAAATCTGCTGCTGGTCCTTTTATGGTCAGGGCCGGAAGATTGCTG
AAGTCTGTTGTACCTCCATCGTCTATGCCACTGAGAAGAAACAGACCAAGGTGGAGTTTCCCGAAGCCCG
GATTTATGAGGAGACCCTGAACATTTTGCTGTATGAGGCCCAGGATGGCCGGGGACCTGACAATGCGCTC
CTGGAGGCCACAGGCGGGGCGGCGGGGCGCTCCCACCACCTGGACGAGGACGAGGAGCGGGAGCGGATCG
AGCGCGTGCGGAGGATCCACATCAAGCGCCCTGATGACCGGGCCCACCTTCACCAGTGAGCAGGCAAGAG
ACCGAGCCGGCCTCCTCTCACCGCCCCCACTCCCTGCCGTGCTACACCCAGATCCTGTGCAGGCTGCCGG
GCCCCTTCTGCTTCCCTTGGAGCCTGGAGATACTTTTGTAACAAGCCAGATGATTATTTTGGTATTGCTT
GACAAGGCAAATTGATTGTCTTGACCCAGGCGTATGACCCCTGTCGTTGAACAAGCTGTGTCTAAGATCT
CTACTTTTCATGAGAATCTGAGACTCTTTGGAGCCAGGCTTTCTCGGTTCTCAGAGGAAAAGTATGAATG
AGTGTGAAGTGTATGTGAGAACTTTTGTTTGCAATATTTATTTTTGTGGGTGTCGACTTCCTATGTGGGC
TTTTTGGGTGACACTCCCTTAAGGGTTCAGTTTGACAATTCTGAGAGTTGTCCTGCAGTTGGAGGCCACC
AGAGGTATCTGAGCTCCCTGCTTCCTATTTCATAATCCTCCAGCCCCAGCAGGTCCACTCCTGGTTCCTG
TGTGTTTGGCCCGGGCACAATCCCCACTGCTTTGCTAGACGTGCTTTCTGCCATGTGGCTTTGGGCCTAG
AGCTTGTTGATAATTGCAGCTTGTGGCAGTGGAAATATGGCTGAATGAGCGTCTAAATCGTTGAGACCAG
TGCAACTTTGGGTGCAAGGCTTTGTTTAGGGATCAAGCCTTTTGCCACCTTGGGCTGGTCTTTGGCCTGG
TGCTCACTGGGACCCCATATGTCTGCGTAGGAGCAGAACTTTCCATGGCAGTAAGTGTCCAGCTCTGTTT
CTGGTTCTTTCCCCAACTCCAGCCCCGTCCAGTTGTTCTCCTGATTGACCCGACTCCACTCCAGGAAGGC
CATCTGACCCTGTGACAGGCATAGCTCATAAACTACCCCTCCCTGGGATCCCGCTCCTCTTCAGCCTCCT
TCCCCATGAAGCTGGGCTAACTTTCTAAGTCATTTGCTTAGAAATTCAGTGTGGCCCATACCCTTTGTC
CTCCCAGCCTGGCATCCAGGCAGGGACACCCTCACACCACCAGCCCCAGGGAGCTTCCCTGCTATAAACA
CAGACCCCCTTGTCTTTGCCTCTGATTTTTACACAGTGTAGAGTGGCCAGCAGTGAACAGGTTGAGGATG
TGCGGGTAGATAGATAACTTTGGGTCTGGTTTGTGTCTGTGTTCATGTTTGTTTAAGGGATATGTGTGAC
TGTGGGTGGGGACGTGTGCTTGTGGGGCACAGGTGGCGGCCCCTGCTGGAGCCCGGCTGGGCGCAGCGCC
TATGTAGGACGGGTGTTCTCAGTGACCTACCTCCCAGGCTCCTCTGCACCTGCAAAGGAACAGGAGTGAG
TCGTGACTGACAGGGGTGGTTGAGACTAGACTAGGTAGAGTAGTTACCAGGAGATGTGAATGTGCGTCAG
GTGATGGATGGGTTTGTCAAGGGAATCGTTACCGTTTTATACCAAAGGTATTAACATGGGCAGCCTTTGA
CACATGTATTCCAAAAACGAGTTTATATTTTCAAACGGTTTTTACAGCTTAGACTTTGTACTTACTGCCC
TGCCTGTGACAGTTGTATGCCTTCATTTGTATCCAACAGCAAAGTCTACAATAAAACTTTAAAACAATC
ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human MSTP028 mRNA sequence - var3 (public gi: 25303941) (SEQ ID NO: 92)
CCGGGTTTGGAGACTCCTGCGTCCTCCGACTTTTCATGGAAGAGATGTCAGGAGAAAGTGTGGTGAGCTC
AGCGGTGCCAGCGGCTGCTACCCGCACCACTTCCTTCAAGGGCACGAGCCCCAGCTCCAAATACGTGAAG
CTGAATGTGGGTGGAGCCCTCTACTATACCACCATGCAGACGCTGACCAAGCAGGACACCATGCTGAAGG
CCATGTTCAGCGGGCGCATGGAAGTGCTCACCGACAGTGAAGGCTGGATCCTCATTGACCGCTGTGGGAA
GCACTTTGGTACGATACTCAACTACCTTCGAGACGGGGCGGTGCCTTTACCCGAGAGCCGCCGGGAGATC
GAGGAGCTGCTAGCAGAAGCCAAGTACTACCTAGTCCAAGGCCTGGTGGAAGAGTGCCAGGCGGCCCTAC
AAAACAAAGATACTTATGAGCCTTTCTGCAAGGTCCCTGTGATCACCTCATCCAAGGAAGAACAAAAACT
TATAGCGACTTCAAATAAGCCAGCCGTGAAGTTGCTCTACAACAGAAGTAACAACAAATACTCATATACC
AGCAATTCTGACGACAATATGTTGAAAAACATTGAACTGTTTGATAAGCTGTCTCTGCGCTTTAACGGAA

GGGTCCTGTTCATAAAGGATGTCATTGGGGATGAAATCTGCTGCTGGTCCTTTTATGGTCAGGGCCGGAA
GATTGCTGAAGTCTGTTGTACCTCCATCGTCTATGCCACTGAGAAGAAACAGACCAAGGTGGAGTTTCCC
GAAGCCCGGATTTATGAGGAGACCCTGAACATTTTGCTGTATGAGGCCCAGGATGGCCGGGGACCTGACA
ATGCGCTCCTGGAGGCCACAGGCGGGGCGGCGGGGCGCTCCCACCACCTGGACGAGGACGAGGAGCGGGA
GCGGATCGAGCGCGTGCGGAGGATCCACATCAAGCGCCCTGATGACCGGGCCCACCTCCACCAGTGAGCA
GGCAAGAGACCGAGCCGCCCTCCTCTCACCGCCCCCACTCCCTGCCGTGCTACACCCAGATCCTGTGCAG
GCTGCCGGGCCCCTTCTGCTTCCCTTGGAGCCTGGAGATACTTTTGTAACAAGCCAGATGATTATTTTGG
TATTGCTTGACAAGGCAAATTGATTGTCTTGACCCAGGCGTATGACCCCTGTCGTTGAACAAGCTGTGTC
TAAGATCTCTACTTTTCATGAGAATCTGAGACTCTTTGGAGCCAGGCTTTCTCGGTTCTCAGAGGAAAAG
TATGAATGAGTGTGAAGTGTATGTGAGAACTTTTGTTTGCAATATTTATTTTTGTGGGTGTCGACTTCCT
ATGTGGGCTTTTTGGGTGACACTCCCTTAAGGGTTCAGTTTGACAATTCTGAGAGTTGTCCTGCAGTTGG
AGGCCACCAGAGGTATCTGAGCTCCCTGCTTCCTATTTCATAATCCTCCAGCCCCAGCAGGTCCACTCCT
GGTTCCTGTGTGTTTGGCCCGGGCACAATCCCCACTGCTTTGCTAGACGTGCTTTCTGCCATGTGGCTTT
GGGCCTAGAGCTTGTTGATAATTGCAGCTTGTGGCAGTGGAAATATGGCTGAATGAGTGTCTAAATCGTT
GAGACCAGTGCAACTTTGGGTGCAAGGCTTTGTTTAGGGATCAAGCCTTTTGCCACCTTGGGCTGGTCTT
TGGCCTGGTGCTCACTGGGACCCCATATGTCTGCGTAGGAGCAGAACTTTCCATGGCAGTAAGTGTCCAG
CTCTGTTTCTGGTTCTTTCCCCAACTCCAGCCCCGTCCAGTTGTTCTCCTGATTGACCCGACTCCACTCC
AGGAAGGCCATCTGACCCTGTGACAGGCATAGCTCATAAACTACCCCTCCCTGGGATCCCGCTCCTCTTC
AGCCTCCTTCCCCATGAAGCTGGGCTAACTTTCTAAGTCATTTTGCTTAGAAATTCAGTGTGGCCCATAC
CCTTTGTCCTCCCAGCCTGGCATCCAGGCAGGGACACCCTCACACCACCAGCCCCAGGGAGCTTCCCTGC
TATAAACACAGACCCCCTTGTCTTTGCCTCTGATTTTTACACAGTGTAGAGTGGCCAGCAGTGAACAGGT
TGAGGATGTGCGGGTAGATAGATAACTTTGGGTCTGGTTTGTGTCTGTGTTCATGTTTGTTTAAGGGATG
TGTGACTGTGGGTGGGGACGTGTGCTTGTGGGGCACAGGTGGCCCCTGCTGGAGCCCGGCTGGGCGCAGC
GCCTATGTAGGACGGGTGTTCTCAGTGACCTACCTCCCAGGCTCCTCTGCACCTGCAAAGGAACAGGAGT
GAGTCGTGACTGACAGGGGTGGTTGAGACTAGACTAGGTAGAGTAGTTACCAGGAGATGTGAATGTGCGT
CAGGTGATGGATGGGTTTGTCAAGGGAATCGTTACCGTTTTATACCAAAGGTATTAACATGGGCAGCCTT
TGACACATGTATTCCAAAAACGAGTTTATATTTTCAAACGGTTTTTACAGCTTAGACTTTGTACTTACTG
CCCTGCCTGTGACAGTTGTATGCCTTCATTTTGTATCCAACAGCAAAGTCTACAATAAAACTTTAAAACA
ATCATGACTGAATGTCAAAATCGTGTATTGGGCAGATGCTTTTTAAACTGTCGTGTGAGAAACTTTTATA
TTAGGCCATTTGGATTTTATTAAGTGCTAAGGAAAGAGGGCTTACAAAATGTTTCGTAAATATTTTATAC
TGTTTAAGTGTTAAACACCAACCCTGTCTTTCTTTTGGGTTGAGCTTTTTTAGAAAGTCGAAGTGAATGT
TGGCCAGGAAAATGGAAAAGCCATTGTATAAATTTTTTTTTGAGGCGGAGTCTTGCTCTATTGGCCAGGC
TGGAGTGTAGTGGCACCATCTCCACTTACCACAACTTGTGCCTCCTGGGTTCAAGCGATTCTGCTGCCTC
AGCCTCCCGAGTAGCTGGGATTGCAGGTACCCATCAGCCCATGCCCAGCTAATTTTGTATTTTTAGTAGA
GATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCCTGTGATCCGACCACCTTGGCCTC
CCAAAGTGCTGGGATTACAGGTGTGAGTCACCACACCTGGCTGCATAGTGTTTTAAATGTTTGTGTGAAG
AATGAGTTTGTGGAACAATTTGATTTGCTGTGGCCTCTATGCCTAATGAGCTAGTGTTTCTGGCAGCTCT
CTCTACCCAACTTTGCACTTGTAGTTTTGAGTCTTTGTCTCTCTGGAATATGAACAGGTTTATAAAACAT
TCCATGGTGAACAATTCTGTCGGCTGCATTATAGCCATGAGTGAATAGACAGCATTGGCTGGTCCAAGCT
CTGTTATTGAGTATACAAGGAACTGATTTTTCTTATGTTAGCACTAAGGGCAAAAACCAATATTTATAAT
GTAAGCACTATCCAGGTAAAACACTGGCCCAAGATTTGGTAAAGAGATTTCATTGCAATGTAATAACTAC
AGTTTTTACAAATTGGAACAGCTTTGGTGTGTCGTAATCAAGGGTTTTTTTGTTTGTTTGTTTCAAAT
AAGCCATCTGATTGTGGTGACTGGGGCCCATGTCCAAGACAATTCCTGGCATATTCTGTCACCCTCCCGT
GGGGCGATCACTGTGTGGGGACCCCATTCCCCAGTTAAAGTGTGTCTCTGTACCTTACAACAGCGATTCA
GGACCCAAGTGTGAACAACACTCAGCCCGCCCTCTGGAGCGTGTGCTGTCTTTAGGGCTCTACCCAAAGT
CACTGTAACAGTTAAGTGTGTCATTAACCTTTCTGTCTCTTTGCGCCATAAAAAAATGCTCAAAGTTTTA
GATGTAGCCACTGTATGTTGTACAAACGTTGGCGACATGTAAAATAAAAGTCATAAAATGCAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human MSTP028 mRNA sequence - var4 (public gi: 16552440) (SEQ ID NO: 93)
AGTCCGGGTTTGGAGACTCCTGCGTCCTCCGACTTTTCATGGGCCCTGACATGGCAGGTGATATCCAGGA
CACTGTTGGGTGCCATGGAGTTGGGGAGAGTTGGCCAGAAGAGTTGGATAACCTTGAATTGAATATTGTC
CATGACTGTCGCGTTGCTTCTGCTGTTGCAAGCTGCCTCCCTTTACTCCCAGTCCCATTTACAAAAATAA
CGCTTGTGTTTTACCAGTTATAGTTGTAGTACCCATTCATTATAGAAAATCTGGAAAAGCTAGACAATTC
TTTTTCAGTTTCAGGGAATAGTTCAAACAAGTTATGTGCTGTCAGTGCCCTGCAGCCAAAAAGCACGAGG
AGCATACCTGTAGTCAAGCAAAGTTGGGTTTATTTCCTTGTTGCATTGGGGTGGGGAAGAACTGTGGGAC
ATCTCAGAGAAGGGCTGTGGGCTTGTGTTGGGTGATTTGAGAGACAGTTCAGAGAAGTGGGGCTTTGCTC
TGTGTTGGATGCTGCTGGGAAGCAGGGCTAATTCTGTGATTGGGTCTCAGTGATTCCTGACTTGAAAGCA
GGAAGAATGGAAGGAGGCTAAACTTCTCATTGGTAAAGCAGCAGCTGTAACTCCTATTAGCCAGGATAGG
GGATCTTTGGTCATTTTTGTATTTTGGATAATGTTTATGTTTTTGTCTGTGTCCGGACATGATGACTGAA
TGGTCCTGTTTTTGTCTTGTCGCAAGGGCACAGAGTGGCCTTGTCTGAGGGTGATGTGCTGTGAAAAACT
GTTGATGTTCAATGGAATGGTAGGGCCAGCCGTGGGGGCTACCCCAGATTCAGCAAAGATTCTGCCCAC
CCTTGCACATTTCCACCTCTACAGTTTTACCTGTTCATTTCAGACATGTTTGCTCTGAGTACACATGTGC

CGGATACCAATCTCACTTTCCAGGCCTGCGTAAATCAGCCACTGTATCCATTTCTTTGAGATGTACAGAG
AGTCAGCCATGCTATCAGGGAGATGGTAGTGGGATCTTGCTCTTTTGGGCAGCACTAGTCTAGGAGGTCT
AATTTTGCAATAACTTGGTTCCAAAAGTTTCCATGTCTGTTGTTTAGTCCTCAGAAACACCTTCTCCCTA
CAGGAAGTGATAGGAGTGCGAGCTGGAATCCCATTCAACTTCATAAAGCTTATTTCATCTGTGATGCAGC
TGAAAAATGACACTTAGCTAGCTATTGAGTGGTACATGGCAATAAGGAAATGTAAAGAGACCTGGGCAGT
GCTTTAGGCTGTTTTAGGGTGCAGCCAGGGTGTTCATGTATACAGGTGCTAGGCAGAAAGGAAGTGCTTA
TAACACAAGAGTTAGGGGCACCCTTGTGCCTGCAGGGTCGACAGGCAGGGTCAGTGTATGAGGCTTTTTG
GGTGGGTCTTGGGACAAACTAGGGGATGCATGGCCCTCTCTAGGGGTCATCCAATACCCCAGCTCTGACC
AGTTGTTCCCCTGCTAGCCCAGTTGGCCTCTGATTTTAGGAGAAGCCAGAAGTCCAGATTTTTCTGTGAG
CTCTCCTTAGTTGACCACATTGGAAGCAAACTTTTAAATGCTGTGTATGCGTGGCCCAAGCAAAACACAT
CTGGAGGCCAGATTGAATCCACAGGCTGAAAGCAGTCAACCAGGCCTGATGTCATGACCCTGTATCCTCT
CCACTGGCAGGAAGAGATGTCAGGAGAAAGTGTGGTGAGCTCAGCGGTGCCAGCGGCTGCTACCCGCACC
ACTTCCTTCAAGGGCACGAGCCCCAGCTCCAAATACGTGAAGCTGAATGTGGGTGGAGCCCTCTACTATA
CCACCATGCAGACGCTGACCAAGCAGGACACCATGCTGAAGGCCATGTTCAGCGGGCGCATGGAAGTGCT
CACCGACAGTGAAGGCTGGATCCTCATTGACCGCTGTGGGAAGCACTTTGGTACGATACTCAACTACCTT
CGAGACGGGGCGGTGCCTTTACCCGAGAGCCGCCGGGAGATCGAGGAGCTGCTAGCAGAAGCCAAGTACT
ACCTAGTCCAAGGCCTGGTGGAAGAGTGCCAGGCGGCCCTACAACAGAACAAAGATACTTATGAGCCTTT
CTGCAAGGTCCCTGTGATCACCTCATCCAAGGAAGAACAAAAACTTATAGCGACTTCAAATAAGCCAGCC
GTGAAGTTGCTCTACAACAGAAGTAACAACAAATACTCATATACCAGCAATTCTGACGACAATATGTTGA
AAAACATTGAACTGTTTGATAAGCTGTCTCTGCGCTTTAACGGAAGGGTCCTGTTCATAAAGGATGTCAT
TGGGGATGAAATCTGCTGCTGGTCCTTTTATGGTCAGGGCCGGAAGATTGCTGAAGTCTGTTGTACCTCC
ATCGTCTATGCCACTGAGAAGAAACAGACCAAGGTGGAGTTTCCCGAAGCCCGGATTTATGAGGAGACCC
TGAACATTTTGCTGTATGAGGCCCAGGATGGCCGGGGACCTGACAATGCGCTCCTGGAGGCCACAGGCGG
GGCGGCGGGGCGCTCCCACCACCTGGACGAGGACGAGGAGCGGGAGCGGATCGAGCGCGTGCGGAGGATC
CACATCAAGCGCCCTGATGACCGGGCCCACCTCCACCAGTGAGCAGGCAAGAGACCGAGCCGCCCTCCTC
TCACCGCCCCCACTCCCTGCCGTGCTACACCCAGATCCTGTGCAGGCTGCCGGGCCCCTTCTGCTTCCCT
TGGAGCCTGGAGATACTTTTGTAACAAGCCAGATGATTATTTTGGTATTGCTTGACAAGGCAAATTGATT
GTCTTGACCCAGGCGTATGACCCCTGTCGTTGAACAAGCTGTGTCTAAGATCTCTACTTTTCATGAGAAT
CTGAGACTCTTTGGAGCCAGGCTTTCTCGGTTCTCAGAGGAAAAGTATGAATGAGTGTGAAGTGTATGTG

Human MSTP028 mRNA sequence - var5 (public gi: 21750697) (SEQ ID NO: 94)
GCTGGCGTGAGCTGGGTGTTTCCTGCCTCTCTCAGTCCGGGTTTGGAGACTCCTGCGTCCTCCGACTTTT
CATGGAAGAGATGTCAGGAGAAAGTGTGGTGAGCTCAGCGGTGCCAGCGGCTGCTACCCGCACCACTTCC
TTCAAGGGCACGAGCCCCAGCTCCAAATACGTGAAGCTGAATGTGGGTGGAGCCCTCTACTATACCACCA
TGCAGACGCTGACCAAGCAGGACACCATGCTGAAGGCCATGTCCAGCGGGCGCATGGAAGTGCTCACCGA
CAGTGAAGAACAAAGATACTTATGAGCCTTTCTGCAAGGTCCCTGTGATCACCTCATCCAAGGAAGAACA
AAAACTTATAGCGACTTCAAATAAGCCAGCCGTGAAGTTGCTCTACAACAGAAGTAACAACAAATACTCA
TATACCAGCGATTCTGACGACAATATGTTGAAAAACATTGAACTGTTTGATAAGCTGTCTCTGCGCTTTA
ACGGAAGGGTCCTGTTCATAAAGGATGTTATTGGGGATGAAATCTGCTGCTGGTCCTTTTATGGTCAGGG
CCGGAAGATTGCTGAAGTCTGTTGTACCTCCATCGTCTATGCCACTGAGAAGAAACAGACCAAGGTGGAG
TTTCCCGAAGCCCGGATTTATGAGGAGACCCTGAACATTTTGCTGTATGAGGCCCAGGGTGGCCGGGGAC
CTGACAATGCGCTCCTGGAGGCCACAGGCGGGGCGGCGGGGCGCTCCCACCACCTGGACGAGGACGAGGA
GCGGGAGCGGATCGAGCGCGTGCGGAGGATCCACATCAAGCGCCCTGATGACCGGGCCCACCTCCACCAG
TGAGCAGGCAAGAGACCGAGCCGCCCTCCTCTCACCGCCCCCACTCCCTGCCGTGCTACACCCAGATCCT
GTGCAGGCTGCCGGGCCCCTTCTGCTTCCCTTGGAGCCTGGAGATACTTTTGTAACAAGCCAGATGATTA
TTTTGGTATTGCTTGACAAGGCAAATTGATTGTCTTGACCCAGGCGTATGACCCCTGTCGTTGAACAAGC
TGTGTCTAAGATCTCTACTTTTCATGAGAATCTGAGACTCTTTGGAGCCAGGCTTTCTCGGTTCTCAGAG
GAAAAGTATGAATGAGTGTGAAGTGTATGTGAGAACTTTTGTTTGCAATATTTATTTTTGTGGGTGTCGA
CTTCCTGTGTGGGCTTTTTGGGTGACACTCCCTTAAGGGTTCAGTTTGACAATTCTGAGAGTTGTCCTGC
AGTTGGAGGCCACCAGAGGTATCTGAGCTCCCTGCTTCCTATTTCATAATCCTCCAGCCCCAGCAGGTCC
ACTCCTGGTTCCTGTGTGTTTGGCCCGGGCACAATCCCCACTGCTTTGCTAGACGTGCTTTCTGCCATGT
GGCTTTGGGCCTAGAGCTTGTTGATAATTGCAGCTTGTGGCAGTGGAAATATGGCTGAATGAGCGTCTAA
ATCGTTGAGACCAGTGCAACTTTGGGTGCAAGGCTTTGTTTAGGGATCAAGCCTTTTGCCACCTTGGGCT
GGTCTTTGGCCTGGTGCTCACTGGGACCCCATATGTCTGCGTAGGAGCAGAACTTTCCATGGCAGTAAGT
GTCCAGCTCTGTTTCTGGTTCTTTCCCCAACTCCAGCCCCGTCCAGTTGTTCTCCTGATTGACCCGACTC
CACTCCAGGAAGGCCATCTGACCCTGTGACAGGCATAGCTCATAAACTACCCCTCCCTGGGATCCCGCTC
CTCTTCAGCCTCCTTCCCCATGAAGCTGGGCTAACTTTCTAAGTCATTTTGCTTAGAAATTCAGTGTGGC
CCATACCCTTTGTCCTCCCAGCCTGGCATCCAGGCAGGGACACCCTCACACCACCAGCCCCAGGGAGCTT
CCCTGCTATAAACACAGACCCCCTTGTCTTTGCCTCTGATTTTTACACAGTGTAGAGTGGCCAGCAGTGA
ACAGGTTGAGGATGTGCGGGTAGATAGATAACTTTGGGTCTGGTTTGTGTCTGTGTTCATGTTTGTTTAA
GGATATGTGTGACTGTGGGTGGGGACGTGTGCTTGTGGGGCACAGGTGGCGGCCCCTGCTGGAGCCTGG
CTGGGCGCAGCGCCTATGTAGGACGGGTGTTCTCAGTGACCTACCTCCCAGGCTCCTCTGCACCTGCAAA
GGAACAGGAGTGAGTCGTGACTGACAGGGGTGGTTGAGACTAGACTAGGTAGAGTAGTTACCAGGAGATG

TGAATGTGCGTCAGGTGATGGATGGGTTTGTCAAGGGAATCGTTACCGTTTTATACCAAAGGTATTAACA
TGGGCAGCCTTTGACACATGTATTCCAAAAACGAGTTTATATTTCAAACGGTTTTACAGCTTAGACTT
TGTACTTACTGCCCTGCCTGTGACAGTTGTATGCCTTCATTTGTATCCAACAGCAAAGTCTACAATAAA
ACTTTAAAACAATCATG

Human MSTP028 Protein sequence - var1 (public gi: 13994353) (SEQ ID NO: 255)
MEEMSGESVVSSAVPAAATRTTSFKGTSPSSKYVKLNVGGALYYTTMQTLTKQDTMLKAMFSGRMEVLTD
SEGWILIDRCGKHFGTILNYLRDGAVPLPESRREIEELLAEAKYYLVQGLVEECQAALQNKDTYEPFCKV
PVITSSKEEQKLIATSNKPAVKLLYNRSNNKYSYTSNSDDNMLKNIELFDKLSLRFNGRVLFIKDVIGDE
ICCWSFYGQGRKIAEVCCTSIVYATEKKQTKVEFPEARIYEETLNILLYEAQDGRGPDNALLEATGGAAG
RSHHLDEDEERERIERVRRIHIKRPDDRAHLHQ

Human MSTP028 Protein sequence - var2 (public gi: 14042295) (SEQ ID NO: 256)
MSGESVVSSAVPAAATRTTSFKGTSPSSKYVKLNVGGALYYTTMQTLTKQDTMLKAMFSGRMEVLTDSEG
WILIDRCGKHFGTILNYLRDGAVPLPESRREIEELLAEAKYYLVQGLVEECQAALQNKDTYEPFCKVPVI
TSSKEEQKLIATSNKPAVKLLYNRSNNKYSYTSNSDDNMLKNIELFDKLSLRFNGRVLFIKDVIGDEICC
WSFYGQGRKIAEVCCTSIVYATEKKQTKVEFPEARIYEETLNILLYEAQDGRGPDNALLEATGGAAGRSH
HLDEDEERERIERVRRIHIKRPDDRAHLHQ

Unigene Name: PACS-1 **Unigene ID:** Hs.58589

Human PACS-1 mRNA sequence - var1 (public gi: 27781345) (SEQ ID NO: 95)
AGCACGAGTCTGGTTGTGCCGGAGAAAGTCAAAACTCCCATGAAGTCCAGTAAAACGGATCTCCAGGGCT
CTGCCTCCCCCAGCAAAGTGGAGGGGGTGCACACACCCCGGCAGAAGAGGAGCACGCCCCTGAAGGAGCG
GCAGCTCTCCAAGCCCCTAAGTGAGAGGACCAACAGTTCCGACAGCGAGCGCTCCCCAGATCTGGGCCAC
AGCACGCAGATTCCAAGAAAGGTGGTGTATGACCAGCTCAATCAGATCCTGGTGTCAGATGCAGCCCTCC
CAGAAAATGTCATTCTGGTGAACACCACTGACTGGCAGGGCCAGTATGTGGCTGAGCTGCTCCAGGACCA
GCGGAAGCCTGTGGTGTGCACCTGCTCCACCGTGGAGGTCCAGGCCGTGCTGTCCGCCCTGCTCACCCGG
ATCCAGCGCTACTGCAACTGCAACTCTTCCATGCCGAGGCCAGTGAAGGTGGCTGCTGTGGGAGGCCAGA
GCTACCTGAGCTCCATCCTCAGGTTCTTTGTCAAGTCCCTGGCCAACAAGACCTCCGACTGGCTTGGCTA
CATGCGCTTCCTCATCATCCCCCTCGGTTCTCACCCTGTGGCCAAATACTTGGGGTCAGTCGACAGTAAA
TACAGTAGTTCCTTCCTGGATTCTGGTTGGAGAGATCTGTTCAGTCGCTCGGAGCCACCAGTGTCAGAGC
AACTGGACGTGGCAGGGCGGGTGATGCAGTACGTCAACGGGGCAGCCACGACACACCAGCTTCCCGTGGC
CGAAGCCATGCTGACTTGCCGGCATAAGTTCCCTGATGAAGACTCCTATCAGAAGTTTATTCCCTTCATT
GGCGTGGTGAAGGTGGGTCTGGTTGAAGACTCTCCCTCCACAGCAGGCGATGGGGACGATTCTCCTGTGG
TCAGCCTTACTGTGCCCTCCACATCACCACCCTCCAGCTCGGGCCTGAGCCGAGACGCCACGGCCACCCC
TCCCTCCTCCCCATCTATGAGCAGCGCCCTGGCCATCGTGGGGAGCCCTAATAGCCCATATGGGGACGTG
ATTGGCCTCCAGGTGGACTACTGGCTGGGCCACCCCGGGGAGCGGAGGAGGGAAGGCGACAAGAGGGACG
CCAGCTCGAAGAACACCCTCAAGAGTGTCTTCCGCTCAGTGCAGGTGTCCCGCCTGCCCCATAGTGGGGA
GGCCCAGCTTTCTGGCACCATGGCCATGACTGTGGTCACCAAAGAAAAGAACAAGAAAGTTCCCACCATC
TTCCTGAGCAAGAAACCCCGAGAAAAGGAGGTGGATTCTAAGAGCCAGGTCATTGAAGGCATCAGCCGCC
TCATCTGCTCAGCCAAGCAGCAGCAGACTATGCTGAGAGTGTCCATCGATGGGGTCGAGTGGAGTGACAT
CAAGTTCTTCCAGCTGGCAGCCCAGTGGCCCACCCATGTCAAGCACTTTCCAGTGGGACTCTTCAGTGGC
AGCAAGGCCACCTGAGGCCCTGTCTCCCAGCCACTTTCCCTCCTGGCACTGCCACCAGCCTCACCGCCTG
CGGGCAGGGGGAGGCCAGCAGGCCCGGGCCCAGCACCCCTTCCCTGGCACCAGGGTCTGCCTCTCACTCG
CCCAGGTCCCGAAGGACACTGCCACAGGGACGCCTTCCCTCCCCTCCCCTCCAGCCCACCCCTGCACAGC
CCCTCCTCCTTCCCGCTTTTCCCCTTCTCCCTCCTGCTCCAGGCCCAAGGCGTGTTGGTTTTGCCTTCTG
GTGCCCATAGTCCCCTGGACTGAGTCCCCCAGGCCTTCCTTCACCCGACTTCCAAACTCTTCCTTGTGGT
ATCAGTTTCCTTCTCGGAAATGAGAAAGCTGGAATCCTGGTCCCCAGCAGGAGAGCCTAGTCCTCCCCCA
GCCCCTCCAGCCACCAGGGTGTCCTCTAGGATGCAGCTGCCAGATCCACTCACTCTGCTGCCTCCAGCAG
GACCCAAGGCCACTTTCAACTCTTATGGGGTTCTCCACCTGCCCCAGAGCTTCTCAAGGGAGGGTAAGGG
GGCACCCTGAGCCCACAGGACCCCTACTTCACAGCTCACAGGGGCAGGAGGCAGCTCCCCTGCCTCCAGG
ACCCTGTTGCTATGGTGACACAGCGTTTCTAGGACAGAGGGGCCTCCCAGTCTCCCCCCACCACCCGTGC
ACGACTTCCTCACCACCCCCAGGTTCCCTGCAGATGTCGTGTGTGTCCTGAGTGTTTCTTTGGTTCTTTG
CACGCCAAGTCTCTTGGTTGTACCATGTGACACACCCTGTGCACTGGTCGCTGTCTTCGTGGCTTCCACC
CTTGTTAATGATGCTCCTGCCTCTGCCTCCCAGCCCCTCACCCAGCACAGCTCTGCCTGGACTTGGAGAG
ATGGGAGGCAGACCCCCACCACCATACATGCTGTCTGTGGCCCCTCAGACATTCTGTTTCATCTCCCATT
CATCTCCCTCCTCCCACCGTGTCAGTTTTTCTGCCTTTCCCTGCTCTGTTCTTCCCCCTCCTTAGGCCCC
AGCCTGGGCCCAGACCCATCCTCCCAGCCAGGTTTCCCTCCAGCAGGCTCCTTCCCTCCCTGTCACCTCC
CTCTCACCAACCCGGGGTCTGAGCCCCTCATTCCTGACCGTCCGTGTTCTCAGGAGTGGTTGAGGACACA
GGGCCCCAGCCCAGCCCTCTGCACCCCCCAGCCCGGCCATCTGCGCCCCACAGCCCCTTTGGAGCTTTTC
TCTTGTCCTCTCACTCCTTCCCAGAAGTTTTTGCACAGAACTTCATTTGAAAGTGTTTTTCTCATTCTC

CATACCTCCCCCAAGCTCTCCTCCAGCCCTTCCCAGGGCTCAGCCCTGCTGTCCTGAGCGTCTCCTGGGC
CAGAGAGAGGAGATGGGGGTGGGAGGGACTGAGTTGATGTTGGGTTTTTCATTCAATAAATTGGTGATTT
CTTAAAAAAAAAAAAAAAAA

Human PACS-1 mRNA sequence - var2 (public gi: 30962845) (SEQ ID NO: 96)
CCTCGGCCTCCGTAACCCCCGCCTAGCCGGGCCATGGCGGAACGCGGAGGGGCGGGCGGTGGTCCCGGAG
GCGCCGGGGGCGGCAGCGGCCAGCGGGGATCCGGGGTCGCCCAGTCCCCTCAGCAGCCGCCGCCGCAGCA
GCAGCAGCAGCAGCCGCCGCAGCAGCCGACGCCCCCCAAGCTGGCCCAGGCCACCTCGTCGTCCTCGTCC
ACCTCGGCGGCGGCTGCCTCCTCCTCGTCCTCGTCTACCTCCACCTCCATGGCCGTGGCGGTGGCCTCGG
GCTCCGCGCCTCCCGGTGGCCCGGGGCCAGGCCGCACCCCCGCCCCGGTGCAGATGAACCTGTACGCCAC
CTGGGAGGTGGACCGGAGCTCGTCCAGCTGCGTGCCTAGGCTATTCAGCTTGACCCTGAAGAAACTCGTC
ATGCTAAAAGAAATGGACAAAGATCTTAACTCAGTGGTCATCGCTGTGAAGCTGCAGGGTTCAAAAAGAA
TTCTTCGCTCCAACGAGATCGTCCTTCCAGCTAGTGGACTGGTGGAAACAGAGCTCCAATTAACCTTCTC
CCTTCAGTACCCTCATTTCCTTAAGCGAGATGCCAACAAGCTGCAGATCATGCTGCAAAGGAGAAAACGT
TACAAGAATCGGACCATCTTGGGCTATAAGACCTTGGCCGTGGGACTCATCAACATGGCAGAGGTGATGC
AGCATCCTAATGAAGGCGCACTGGTGCTTGGCCTACACAGCAACGTGAAGGATGTCTCTGTGCCTGTGGC
AGAAATAAAGATCTACTCCCTGTCCAGCCAACCCATTGACCATGAAGGAATCAAATCCAAGCTTTCTGAT
CGTTCTCCTGATATTGACAATTATTCTGAGGAAGAGGAAGAGAGTTTCTCATCAGAACAGGAAGGCAGTG
ATGATCCATTGCATGGGCAGGACTTGTTCTACGAAGACGAAGATCTCCGGAAAGTGAAGAAGACCCGGAG
GAAACTAACCTCAACCTCTGCCATCACAAGGCAACCTAACATCAAACAGAAGTTTGTGGCCCTCCTGAAG
CGGTTTAAAGTTTCAGATGAGGTGGGCTTTGGGCTGGAGCATGTGTCCCGCGAGCAGATCCGGGAAGTGG
AAGAGGACTTGGATGAATTGTATGACAGTCTGGAGATGTACAACCCCAGCGACAGTGGCCCTGAGATGGA
GGAGACAGAAAGCATCCTCAGCACGCCAAAGCCCAAGCTCAAGCCTTTCTTTGAGGGGATGTCGCAGTCC
AGCTCCCAGACGGAGATTGGCAGCCTCAACAGCAAAGGCAGCCTCGGAAAAGACACCACCAGCCCTATGG
AATTGGCTGCTCTAGAAAAAATTAAATCTACTTGGATTAAAAACCAAGATGACAGCTTGACTGAAACAGA
CACTCTGGAAATCACTGACCAGGACATGTTTGGAGATGCCAGCACGAGTCTGGTTGTGCCGGAGAAAGTC
AAAACTCCCATGAAGTCCAGTAAAACGGATCTCCAGGGCTCTGCCTCCCCCAGCAAAGTGGAGGGGGTGC
ACACACCCCGGCAGAAGAGGAGCACGCCCCTGAAGGAGCGGCAGCTCTCCAAGCCCCTAAGTGAGAGGAC
CAACAGTTCCGACAGCGAGCGCTCCCCAGATCTGGGCCACAGCACGCAGATTCCAAGAAAGGTGGTGTAT
GACCAGCTCAATCAGATCCTGGTGTCAGATGCAGCCCTCCCAGAAAATGTCATTCTGGTGAACACCACTG
ACTGGCAGGGCCAGTATGTGGCTGAGCTGCTCCAGGACCAGCGGAAGCCTGTGGTGTGCACCTGCTCCAC
CGTGGAGGTCCAGGCCGTGCTGTCCGCCCTGCTCACCCGGATCCAGCGCTACTGCAACTGCAACTCTTCC
ATGCCGAGGCCAGTGAAGGTGGCTGCTGTGGGAGGCCAGAGCTACCTGAGCTCCATCCTCAGGTTCTTTG
TCAAGTCCCTGGCCAACAAGACCTCCGACTGGCTTGGCTACATGCGCTTCCTCATCATCCCCCTCGGTTC
TCACCCTGTGGCCAAATACTTGGGGTCAGTCGACAGTAAATACAGTAGTTCCTTCCTGGATTCTGGTTGG
AGAGATCTGTTCAGTCGCTCGGAGCCACCAGTGTCAGAGCAACTGGACGTGGCAGGGCGGGTGATGCAGT
ACGTCAACGGGGCAGCCACGACACACCAGCTTCCCGTGGCCGAAGCCATGCTGACTTGCCGGCATAAGTT
CCCTGATGAAGACTCCTATCAGAAGTTTATTCCCTTCATTGGCGTGGTGAAGGTGGGTCTGGTTGAAGAC
TCTCCCTCCACAGCAGGCGATGGGGACGATTCTCCTGTGGTCAGCCTTACTGTGCCCTCCACATCACCAC
CCTCCAGCTCGGGCCTGAGCCGAGACGCCACGGCCACCCCTCCCTCCTCCCCATCTATGAGCAGCGCCCT
GGCCATCGTGGGGAGCCCTAATAGCCCATATGGGGACGTGATTGGCCTCCAGGTGGACTACTGGCTGGGC
CACCCCGGGGAGCGGAGGAGGGAAGGCGACAAGAGGGACGCCAGCTCGAAGAACACCCTCAAGAGTGTCT
TCCGCTCAGTGCAGGTGTCCCGCCTGCCCCATAGTGGGGAGGCCCAGCTTTCTGGCACCATGGCCATGAC
TGTGGTCACCAAAGAAAAGAACAAGAAAGTTCCCACCATCTTCCTGAGCAAGAAACCCCGAGAAAAGGAG
GTGGATTCTAAGAGCCAGGTCATTGAAGGCATCAGCCGCCTCATCTGCTCAGCCAAGCAGCAGCAGACTA
TGCTGAGAGTGTCCATCGATGGGGTCGAGTGGAGTGACATCAAGTTCTTCCAGCTGGCAGCCCAGTGGCC
CACCCATGTCAAGCACTTTCCAGTGGGACTCTTCAGTGGCAGCAAGGCCACCTGAGGCCCTGTCTCCCAG
CCACTTTCCCTCCTGGCACTGCCACCAGCCTCACCGCCTGCGGGCAGGGGGAGGCCAGCAGGCCCGGGCC
CAGCACCCCTTCCCTGGCACCAGGGTCTGCCTCTCACTCGCCCAGGTCCCGAAGGACACTGCCACAGGGA
CGCCTTCCCTCCCCTCCCCTCCAGCCCACCCCTGCACAGCCCCTCCTCCTTCCCGCTTTTCCCCTTCTCC
CTCCTGCTCCAGGCCCAAGGCGTGTTGGTTTTGCCTTCTGGTGCCCATAGTCCCCTGGACTGAGTCCCCC
AGGCCTTCCTTCACCCGACTTCCAAACTCTTCCTTGTGGTATCAGTTTCCTTCTCGGAAATGAGAAAGCT
GGAATCCTGGTCCCCAGCAGGAGAGCCTAGTCCTCCCCCAGCCCCTCCAGCCACCAGGGTGTCCTCTAGG
ATGCAGCTGCCAGATCCACTCACTCTGCTGCCTCCAGCAGGACCCAAGGCCACTTTCAACTCTTATGGGG
TTCTCCACCTGCCCCAGAGCTTCTCAAGGGAGGGTAAGGGGGCACCCTGAGCCCACAGGACCCCTACTTC
ACAGCTCACAGGGGCAGGAGGCAGCTCCCCTGCCTCCAGGACCCTGTTGCTATGGTGACACAGCGTTTCT
AGGACAGAGGGGCCTCCCAGTCTCCCCCCACCACCCGTGCACGACTTCCTCACCACCCCCAGGTTCCCTG
CAGATGTCGTGTGTGTCCTGAGTGTTTCTTTGGTTCTTTGCACGCCAAGTCTCTTGGTTGTACCATGTGA
CACACCCTGTGCACTGGTCGCTGTCTTCGTGGCTTCCACCCTTGTTAATGATGCTCCTGCCTCTGCCTCC
CAGCCCCTCACCCAGCACAGCTCTGCCTGGACTTGGAGAGATGGGAGGCAGACCCCCACCACCATACATG
CTGTCTGTGGCCCCTCAGACATTCTGTTTCATCTCCCATTCATCTCCCTCCTCCCACCGTGTCAGTTTTT
CTGCCTTTCCCTGCTCTGTTCTTCCCCCTCCTTAGGCCCCAGCCTGGGCCCAGACCCATCCTCCCAGCCA
GGTTTCCCTCCAGCAGGCTCCTTCCCTCCCTGTCACCTCCCTCTCACCAACCCGGGGTCTGAGCCCCTCA

TTCCTGACCGTCCGTGTTCTCAGGAGTGGTTGAGGACACAGGGCCCCAGCCCAGCCCTCTGCACCCCCCA
GCCCGGCCATCTGCGCCCCACAGCCCCTTTGGAGCTTTTCTCTTGTCCTCTCACTCCTTCCCAGAAGTTT
TTGCACAGAACTTCATTTTGAAAGTGTTTTTCTCATTCTCCATACCTCCCCCAAGCTCTCCTCCAGCCCT
TCCCAGGGCTCAGCCCTGCTGTCCTGAGCGTCTCCTGGGCCAGAGAGAGGAGATGGGGGTGGAGGGACT
GAGTTGATGTTGGGTTTTTCATTCAATAAATTGGTGATTTCTTACCGACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA

Human PACS-1 mRNA sequence - var3 (public gi: 33243994) (SEQ ID NO: 97)
CAGAAAGCATCCTCAGCACGCCAAAGCCCAAGCTCAAGCCTTTCTTTGAGGGGATGTCGCAGTCCAGCTC
CCAGACGGAGATTGGCAGCCTCAACAGCAAAGGCAGCCTCGGAAAAGACACCACCAGCCCTATGGAATTG
GCTGCTCTAGAAAAAATTAAATCTACTTGGATTAAAAACCAAGATGACAGCTTGACTGAAACAGACACTC
TGGAAATCACTGACCAGGACATGTTTGGAGATGCCAGCACGAGTCTGGTTGTGCCGGAGAAAGTCAAAAC
TCCCATGAAGTCCAGTAAAACGGATCTCCAGGGCTCTGCCTCCCCCAGCAAAGTGGAGGGGGTGCACACA
CCCCGGCAGAAGAGGAGCACGCCCCTGAAGGAGCGGCAGCTCTCCAAGCCCCTAAGTGAGAGGACCAACA
GTTCCGACAGCGAGCGCTCCCCAGATCTGGGCCACAGCACGCAGATTCCAAGAAAGGTGGTGTATGACCA
GCTCAATCAGATCCTGGTGTCAGATGCAGCCCTCCCAGAAAATGTCATTCTGGTGAACACCACTGACTGG
CAGGGCCAGTATGTGGCTGAGCTGCTCCAGGACCAGCGGAAGCCTGTGGTGTGCACCTGCTCCACCGTGG
AGGTCCAGGCCGTGCTGTCCGCCCTGCTCACCCGGATCCAGCGCTACTGCAACTGCAACTCTTCCATGCC
GAGGCCAGTGAAGGTGGCTGCTGTGGGAGGCCAGAGCTACCTGAGCTCCATCCTCAGGTTCTTTGTCAAG
TCCCTGGCCAACAAGACCTCCGACTGGCTTGGCTACATGCGCTTCCTCATCATCCCCCTCGGTTCTCACC
CTGTGGCCAAATACTTGGGGTCAGTCGACAGTAAATACAGTAGTTCCTTCCTGGATTCTGGTTGGAGAGA
TCTGTTCAGTCGCTCGGAGCCACCAGTGTCAGAGCAACTGGACGTGGCAGGGCGGGTGATGCAGTACGTC
AACGGGGCAGCCACGACACACCAGCTTCCCGTGGCCGAAGCCATGCTGACTTGCCGGCATAAGTTCCCTG
ATGAAGACTCCTATCAGAAGTTTATTCCCTTCATTGGCGTGGTGAAGGTGGGTCTGGTTGAAGACTCTCC
CTCCACAGCAGGCGATGGGGACGATTCTCCTGTGGTCAGCCTTACTGTGCCCTCCACATCACCACCCTCC
AGCTCGGGCCTGAGCCGAGACGCCACGGCCACCCCTCCCTCCTCCCCATCTATGAGCAGCGCCCTGGCCA
TCGTGGGGAGCCCTAATAGCCCATATGGGGACGTGATTGGCCTCCAGGTGGACTACTGGCTGGGCCACCC
CGGGGAGCGGAGGAGGGAAGGCGACAAGAGGGACGCCAGCTCGAAGAACACCCTCAAGAGTGTCTTCCGC
TCAGTGCAGGTGTCCCGCCTGCCCCATAGTGGGGAGGCCCAGCTTTCTGGCACCATGGCCATGACTGTGG
TCACCAAAGAAAAGAACAAGAAAGTTCCCACCATCTTCCTGAGCAAGAAACCCCGAGAAAAGGAGGTGGA
TTCTAAGAGCCAGGTCATTGAAGGCATCAGCCGCCTCATCTGCTCAGCCAAGCAGCAGCAGACTATGCTG
AGAGTGTCCATCGATGGGGTCGAGTGGAGTGACATCAAGTTCTTCCAGCTGGCAGCCCAGTGGCCCACCC
ATGTCAAGCACTTTCCAGTGGGACTCTTCAGTGGCAGCAAGGCCACCTGAGGCCCTGTCTCCCAGCCACT
TTCCCTCCTGGCACTGCCACCAGCCTCACCGCCTGCGGGCAGGGGGAGGCCAGCAGGCCCGGGCCCAGCA
CCCCTTCCCTGGCACCAGGGTCTGCCTCTCACTCGCCCAGGTCCCGAAGGACACTGCCACAGGGACGCCT
TCCCTCCCCTCCCCTCCAGCCCACCCCTGCACAGCCCCTCCTCCTTCCCGCTTTTCCCCTTCTCCCTCCT
GCTCCAGGCCCAAGGCGTGTTGGTTTTGCCTTCTGGTGCCCATAGTCCCCTGGACTGAGTCCCCCAGGCC
TTCCTTCACCCGACTTCCAAACTCTTCCTTGTGGTATCAGTTTCCTTCTTGGAAATGAGAAAGCTGGAAT
CCTGGTCCCCAGCAGGAGAGCCTAGTCCTCCCCCAGCCCCTCCAGCCACCAGGGTGTCCTCTAGGATGCA
GCTGCCAGATCCACTCACTCTGCTGCCTCCAGCAGGACCCAAGGCCACTTTCAACTCTTATGGGGTTCTC
CACCTGCCCCAGAGCTTCCCAAGGGAGGGTAAGGGGGCACCCTGAGCCCACAGGACCCCTACTTCACAGC
TCACAGGGGCAGGAGGCAGCTCCCCTGCCTCCAGGACCCTGTTGCTATGGTGACACAGCGTTTCTAGGAC
AGAGGGGCCTCCCAGTCTCCCCCCACCACCCGTGCACGACTTCCTCACCACCCCCAGGTTCCCTGCAGAT
GTCGTGTGTGTCCTGAGTGTTTCTTTGGTTCTTTGCACGCCAAGTCTCTTGGTTGTACCATGTGACACAC
CCTGTGCACTGGTCGCTGTCTTCGTGGCTTCCACCCTTGTTAATGATGCTCCTGCCTCTGCCTCCCAGCC
CCTCACCCAGCACAGCTCTGCCTGGACTTGGAGAGATGGGAGGCAGACCCCCACCACCATACATGCTGTC
TGTGGCCCCTCAGACATTCTGTTTCATCTCCCATTCATCTCCCTCCTCCCACCGTGTCAGTTTTTCTGCC
TTTCCCTGCTCTGTTCTTCCCCCTCCTTAGGCCCCAGCCTGGGCCCAGACCCATCCTCCCAGCCAGGTTT
CCCTCCAGCAGGCTCCTTCCCTCCCTGTCACCTCCCTCTCACCAACCCGGGGTCTGAGCCCCTCATTCCT
GACCGTCCGTGTTCTCAGGAGTGGTTGAGGACACAGGGCCCCAGCCCAGCCCTCTGCACCCCCCAGCCCG
GCCATCTGCGCCCCACAGCCCCTTTGGAGCTTTTCTCTTGTCCTCTCACTCCTTCCCAGAAGTTTTTGCA
CAGAACTTCATTTTGAAAGTGTTTTTCTCATTCTCTATACCTCCCCCAAGCTCTCCTCCAGCCCTTCCCA
GGGCTCAGCCCTGCTGTCCTGAGCGTCTCCTGGGCCAGAGAGAGGAGATGGGGGTGGAGGGACTGAGTT
GATGTTGGGTTTTTCATTCAATAAATTGGTGATTTCTTACCGACAAAAAAAAAAAAAAA

Human PACS-1 mRNA sequence - var4 (public gi: 34420884) (SEQ ID NO: 98)
CGCCGCCGCCGCCGCGGGGGAAGCCTGGGAGCCAGATCGGCGTCGCCTCGGCCTCCGTAACCCCCGCCTA
GCCGGGCCATGGCGGAACGCGGAGGGGCGGGCGGTGGTCCCGGAGGCGCCGGGGGCGGCAGCGGCCAGCG
GGGATCCGGGGTCGCCCAGTCCCCTCAGCAGCCGATGGCGGAACGCGGAGGGGCGGGCGGTGGTCCCGGA
GGCGCCGGGGGCGGCAGCGGCCAGCGGGGATCCGGGGTCGCCCAGTCCCCTCAGCAGCCGCCGCCGCAGC
AGCAGCAGCAGCAGCCGCCGCAGCAGCCGACGCCCCCCAAGCTGGCCCAGGCCACCTCGTCGTCCTCGTC
CACCTCGGCGGCGGCTGCCTCCTCCTCGTCCTCGTCTACCTCCACCTCCATGGCCGTGGCGGTGGCCTCG

GGCTCCGCGCCTCCCGGTGGCCCGGGGCCAGGCCGCACCCCCGCCCCGGTGCAGATGAACCTGTACGCCA
CCTGGGAGGTGGACCGGAGCTCGTCCAGCTGCGTGCCTAGGCTATTCAGCTTGACCCTGAAGAAACTCGT
CATGCTAAAAGAAATGGACAAAGATCTTAACTCAGTGGTCATCGCTGTGAAGCTGCAGGGTTCAAAAAGA
ATTCTTCGCTCCAACGAGATCGTCCTTCCAGCTAGTGGACTGGTGGAAACAGAGCTCCAATTAACCTTCT
CCCTTCAGTACCCTCATTTCCTTAAGCGAGATGCCAACAAGCTGCAGATCATGCTGCAAAGGAGAAAACG
TTACAAGAATCGGACCATCTTGGGCTATAAGACCTTGGCCGTGGACTCATCAACATGGCAGAGGTGATG
CAGCATCCTAATGAAGGCGCACTGGTGCTTGGCCTACACAGCAACGTGAAGGATGTCTCTGTGCCTGTGG
CAGAAATAAAGATCTACTCCCTGTCCAGCCAACCCATTGACCATGAAGGAATCAAATCCAAGCTTTCTGA
TCGTTCTCCTGATATTGACAATTATTCTGAGGAAGAGGAAGAGAGTTTCTCATCAGAACAGGAAGGCAGT
GATGATCCATTGCATGGGCAGGACTTGTTCTACGAAGACGAAGATCTCCGGAAAGTGAAGAAGACCCGGA
GGAAACTAACCTCAACCTCTGCCATCACAAGGCAACCTAACATCAAACAGAAGTTTGTGGCCCTCCTGAA
GCGGTTTAAAGTTTCAGATGAGGTGGGCTTTGGGCTGGAGCATGTGTCCCGCGAGCAGATCCGGGAAGTG
GAAGAGGACTTGGATGAATTGTATGACAGTCTGGAGATGTACAACCCCAGCGACAGTGGCCCTGAGATGG
AGGAGACAGAAAGCATCCTCAGCACGCCAAAGCCCAAGCTCAAGCCTTTCTTTGAGGGGATGTCGCAGTC
CAGCTCCCAGACGGAGATTGGCAGCCTCAACAGCAAAGGCAGCCTCGGAAAAGACACCACCAGCCCTATG
GAATTGGCTGCTCTAGAAAAAATTAAATCTACTTGGATTAAAAACCAAGATGACAGCTTGACTGAAACAG
ACACTCTGGAAATCACTGACCAGGACATGTTTGGAGATGCCAGCACGAGTCTGGTTGTGCCGGAGAAAGT
CAAAACTCCCATGAAGTCCAGTAAAACGGATCTCCAGGGCTCTGCCTCCCCCAGCAAAGTGGAGGGGGTG
CACACACCCCGGCAGAAGAGGAGCACGCCCCTGAAGGAGCGGCAGCTCTCCAAGCCCCTAAGTGAGAGGA
CCAACAGTTCCGACAGCGAGCGCTCCCCAGATCTGGGCCACAGCACGCAGATTCCAAGAAAGGTGGTGTA
TGACCAGCTCAATCAGATCCTGGTGTCAGATGCAGCCCTCCCAGAAAATGTCATTCTGGTGAACACCACT
GACTGGCAGGGCCAGTATGTGGCTGAGCTGCTCCAGGACCAGCGGAAGCCTGTGGTGTGCACCTGCTCCA
CCGTGGAGGTCCAGGCCGTGCTGTCCGCCCTGCTCACCCGGATCCAGCGCTACTGCAACTGCAACTCTTC
CATGCCGAGGCCAGTGAAGGTGGCTGCTGTGGGAGGCCAGAGCTACCTGAGCTCCATCCTCAGGTTCTTT
GTCAAGTCCCTGGCCAACAAGACCTCCGACTGGCTTGGCTACATGCGCTTCCTCATCATCCCCCTCGGTT
CTCACCCTGTGGCCAAATATTTGGGGTCAGTCGACAGTAAATACAGTAGTTCCTTCCTGGATTCTGGTTG
GAGAGATCTGTTCAGTCGCTCGGAGCCACCAGTGTCAGAGCAACTGGACGTGGCAGGGCGGGTGATGCAG
TACGTCAACGGGGCAGCCACGACACACCAGCTTCCCGTGGCCGAAGCCATGCTGACTTGCCGGCATAAGT
TCCCTGATGAAGACTCCTATCAGAAGTTTATTCCCTTCATTGGCGTGGTGAAGGTGGGTCTGGTTGAAGA
CTCTCCCTCCACAGCAGGCGATGGGGACGATTCTCCTGTGGTCAGCCTTACTGTGCCCTCCACATCACCA
CCCTCCAGCTCGGGCCTGAGCCGAGACGCCACGGCCACCCCTCCCTCCTCCCCATCTATGAACAGCGCCC
TGGCCATCGTGGGGAGCCCTAATAGCCCATATGGGGACGTGATTGGCCTCCAGGTGGACTACTGGCTGGG
CCACCCCGGGGAGCGGAGGAGGGAAGGCGACAAGAGGGACGCCAGCTCGAAGAACACCCTCAAGAGTGTC
TTCCGCTCAGTGCAGGTGTCCCGCCTGCCCCATAGTGGGGAGGCCCAGCTTTCTGGCACCATGGCCATGA
CTGTGGTCACCAAAGAACTGAACAAGAAAGTTCCCACCATCTTCCTGAGCAAGAAACCCCGAGAAAAGGA
GGTGGATTCTAAGAGCCAGGTCATTGAAGGCATCAGCCGCCTCATCTGCTCAGCCAAGCAGCAGCAGACT
ATGCTGAGAGTGTCCATCGATGGGGTCGAGTGGAGTGACATCAAGTTCTTCCAGCTGGCAGCCCAGTGGC
CCACCCATGTCAAGCACTTTCCAGTGGGACTCTTCAGTGGCAGCAAGGCCACCTAG

Human PACS-1 mRNA sequence - var5 (public gi: 6330230) (SEQ ID NO: 99)
CTGCCATCACAAGGCAACCTAACATCAAACAGAAGTTTGTGGCCCTCCTGAAGCGGTTTAAAGTTTCAGA
TGAGGTGGGCTTTGGGCTGGAGCATGTGTCCCGCGAGCAGATCCGGGAAGTGGAAGAGGACTTGGATGAA
TTGTATGACAGTCTGGAGATGTACAACCCCAGCGACAGTGGCCCTGAGATGGAGGAGACAGAAAGCATCC
TCAGCACGCCAAAGCCCAAGCTCAAGCCTTTCTTTGAGGGGATGTCGCAGTCCAGCTCCCAGACGGAGAT
TGGCAGCCTCAACAGCAAAGGCAGCCTCGGAAAAGACACCACCAGCCCTATGGAATTGGCTGCTCTAGAA
AAAATTAAATCTACTTGGATTAAAAACCAAGATGACAGCTTGACTGAAACAGACACTCTGGAAATCACTG
ACCAGGACATGTTTGGAGATGCCAGCACGAGTCTGGTTGTGCCGGAGAAAGTCAAAACTCCCATGAAGTC
CAGTAAAACGGATCTCCAGGGCTCTGCCTCCCCCAGCAAAGTGGAGGGGGTGCACACACCCCGGCAGAAG
AGGAGCACGCCCCTGAAGGAGCGGCAGCTCTCCAAGCCCCTAAGTGAGAGGACCAACAGTTCCGACAGCG
AGCGCTCCCCAGATCTGGGCCACAGCACGCAGATTCCAAGAAAGGTGGTGTATGACCAGCTCAATCAGAT
CCTGGTGTCAGATGCAGCCCTCCCAGAAAATGTCATTCTGGTGAACACCACTGACTGGCAGGGCCAGTAT
GTGGCTGAGCTGCTCCAGGACCAGCGGAAGCCTGTGGTGTGCACCTGCTCCACCGTGGAGGTCCAGGCCG
TGCTGTCCGCCCTGCTCACCCGGATCCAGCGCTACTGCAACTGCAACTCTTCCATGCCGAGGCCAGTGAA
GGTGGCTGCTGTGGGAGGCCAGAGCTACCTGAGCTCCATCCTCAGGTTCTTTGTCAAGTCCCTGGCCAAC
AAGACCTCCGACTGGCTTGGCTACATGCGCTTCCTCATCATCCCCCTCGGTTCTCACCCTGTGGCCAAAT
ACTTGGGGTCAGTCGACAGTAAATACAGTAGTTCCTTCCTGGATTCTGGTTGGAGAGATCTGTTCAGTCG
CTCGGAGCCACCAGTGTCAGAGCAACTGGACGTGGCAGGGCGGGTGATGCAGTACGTCAACGGGGCAGCC
ACGACACACCAGCTTCCCGTGGCCGAAGCCATGCTGACTTGCCGGCATAAGTTCCCTGATGAAGACTCCT
ATCAGAAGTTTATTCCCTTCATTGGCGTGGTGAAGGTGGGTCTGGTTGAAGACTCTCCCTCCACAGCAGG
CGATGGGGACGATTCTCCTGTGGTCAGCCTTACTGTGCCCTCCACATCACCACCCTCCAGCTCGGGCCTG
AGCCGAGACGCCACGGCCACCCCTCCCTCCTCCCCATCTATGAGCAGCGCCCTGGCCATCGTGGGGAGCC
CTAATAGCCCATATGGGGACGTGATTGGCCTCCAGGTGGACTACTGGCTGGGCCACCCCGGGGAGCGGAG
GAGGGAAGGCGACAAGAGGGACGCCAGCTCGAAGAACACCCTCAAGAGTGTCTTCCGCTCAGTGCAGGTG

TCCCGCCTGCCCCATAGTGGGGAGGCCCAGCTTTCTGGCACCATGGCCATGACTGTGGTCACCAAAGAAA
AGAACAAGAAAGTTCCCACCATCTTCCTGAGCAAGAAACCCCGAGAAAAGGAGGTGGATTCTAAGAGCCA
GGTCATTGAAGGCATCAGCCGCCTCATCTGCTCAGCCAAGCAGCAGCAGACTATGCTGAGAGTGTCCATC
GATGGGGTCGAGTGGAGTGACATCAAGTTCTTCCAGCTGGCAGCCCAGTGGCCCACCCATGTCAAGCACT
TTCCAGTGGGACTCTTCAGTGGCAGCAAGGCCACCTGAGGCCCTGTCTCCCAGCCACTTTCCCTCCTGGC
ACTGCCACCAGCCTCACCGCCTGCGGGCAGGGGGAGGCCAGCAGGCCCGGGCCCAGCACCCCTTCCCTGG
CACCAGGGTCTGCCTCTCACTCGCCCAGGTCCCGAAGGACACTGCCACAGGGACGCCTTCCCTCCCCTCC
CCTCCAGCCCACCCCTGCACAGCCCCTCCTCCTTCCCGCTTTTCCCCTTCTCCCTCCTGCTCCAGGCCCA
AGGCGTGTTGGTTTTGCCTTCTGGTGCCCATAGTCCCCTGGACTGAGTCCCCCAGGCCTTCCTTCACCCG
ACTTCCAAACTCTTCCTTGTGGTATCAGTTTCCTTCTCGGAAATGAGAAAGCTGGAATCCTGGTCCCCAG
CAGGAGAGCCTAGTCCTCCCCCAGCCCCTCCAGCCACCAGGGTGTCCTCTAGGATGCAGCTGCCAGATCC
ACTCACTCTGCTGCCTCCAGCAGGACCCAAGGCCACTTTCAACTCTTATGGGGTTCTCCACCTGCCCCAG
AGCTTCCCAAGGGAGGGTAAGGGGGCACCCTGAGCCCACAGGACCCCTACTTCACAGCTCACAGGGGCAG
GAGGCAGCTCCCCTGCCTCCAGGACCCTGTTGCTATGGTGACACAGCGTTTCTAGGACAGAGGGGCCTCC
CAGTCTCCCCCCACCACCCGTGCACGACTTCCTCACCACCCCCAGGTTCCCTGCAGATGTCGTGTGTGTC
CTGAGTGTTTCTTTGGTTCTTTGCACGCCAAGTCTCTTGGTTGTACCATGTGACACACCCTGTGCACTGG
TCGCTGTCTTCGTGGCTTCCACCCTTGTTAATGATGCTCCTGCCTCTGCCTCCCAGCCCCTCACCCAGCA
CAGCTCTGCCTGGACTTGGAGAGATGGGAGGCAGACCCCCACCACCATACATGCTGTCTGTGGCCCCTCA
GACATTCTGTTTCATCTCCCATTCATCTCCCTCCTCCCACCGTGTCAGTTTTCTGCCTTTCCCTGCTCT
GTTCTTCCCCCTCCTTAGGCCCCAGCCTGGGCCCAGACCCATCCTCCCAGCCAGGTTTCCCTCCAGCAGG
CTCCTTCCCTCCCTGTCACCTCCCTCTCACCAACCCGGGGTCTGAGCCCCTCATTCCTGACCGTCCGTGT
TCTCAGGAGTGGTTGAGGACACAGGGCCCCAGCCCAGCCCTCTGCACCCCCCAGCCCGGCCATCTGCGCC
CCACAGCCCCTTTGGAGCTTTTCTCTTGTCCTCTCACTCCTTCCCAGAAGTTTTTGCACAGAACTTCATT
TTGAAAGTGTTTTTCTCATTCTCCATACCTCCCCCAAGCTCTCCTCCAGCCCTTCCCAGGGCTCAGCCCT
GCTGTCCTGAGCGTCTCCTGGGCCAGAGAGAGGAGATGGGGGTGGAGGGACTGAGTTGATGTTGGGTTT
TTCATTCAATAAATTGGTGATTTCTTACCG

Human PACS-1 mRNA sequence - var6 (public gi: 7022110) (SEQ ID NO: 100)
CCCTAAGTGAGAGGACCAACAGTTCCGACAGCGAGCGCTCCCCAGATCTGGGCCACAGCACGCAGATTCC
AAGAAAGGTGGTGTATGACCAGCTCAATCAGATCCTGGTGTCAGATGCAGCCCTCCCAGAAAATGTCATT
CTGGTGAACACCACTGACTGGCAGGGCCAGTATGTGGCTGAGCTGCTCCAGGACCAGCGGAAGCCTGTGG
TGTGCACCTGCTCCACCGTGGAGGTCCAGGCCGTGCTGTCCGCCCTGCTCACCCGGATCCAGCGCTACTG
CAACTGCAACTCTTCCATGCCGAGGCCAGTGAAGGTGGCTGCTGTGGGAGGCCAGAGCTACCTGAGCTCC
ATCCTCAGGTTCTTTGTCAAGTCCCTGGCCAACATGACCTCCGACTGGCTTGGCTACATGCGCTTCCTCA
TCATCCCCCTCGGTTCTCACCCTGTGGCCAAATACTTGGGGTCAGTCGACAGTAAATACAGTAGTTCCTT
CCTGGATTCTGGTTGGAGAGATCTGTTCAGTCGCTCGGAGCCACCAGTGTCAGAGCAACTGGACGTGGCA
GGGCGGGTGATGCAGTACGTCAACGGGGCAGCCACGACACACCAGCTTCCCGTGGCCGAAGCCATGCTGA
CTTGCCGGCATAAGTTCCCTGATGAAGACTCCTATCAGAAGTTTATTCCCTTCATTGGCGTGGTGAAGGT
GGGTCTGGTTGAAGACTCTCCCTCCACAGCAGGCGATGGGGACGATTCTCCTGTGGTCAGCCTTACTGTG
CCCTCCACATCACCACCCTCCAGCTCGGGCCTGAGCCGAGACGCCACGGCCACCCCTCCCTCCTCCCCAT
CTATGAGCAGCGCCCTGGCCATCGTGGGGAGCCCTAATAGCCCATATGGGGACGTGATTGGCCTCCAGGT
GGACTACTGGCTGGGCCACCCCGGGGAGCGGAGGAGGGAAGGCGACAAGAGGGACGCCAGCTCGAAGAAC
ACCCTCAAGAGTGTCTTCCGCTCAGTGCAGGTGTCCCGCCTGCCCCATAGTGGGGAGGCCCAGCTTTCTG
GCACCATGGCCATGACTGTGGTCACCAAAGAAAAGAACAAGAAAGTTCCCACCATCTTCCTGAGCAAGAA
ACCCCGAGAAAAGGAGGTGGATTCTAAGAGCCAGGTCATTGAAGGCATCAGCCGCCTCATCTGTTCTTCC
CCCTCCTTAGGCCCCAGCCTGGGCCCAGACCCATCCTCCCAGCCAGGTTTCCCTCCAGCAGGCTCCTTCC
CTCCCTGTCACCTCCCTCTCACCAACCCGGGGTCTGAGCCCCTCATTCCTGACCGTCCGTGTTCTCAGGA
GTGGTTGAGGACACAGGGCCCCAGCCCAGCCCTCTGCACCCCCCAGCCCGGCCATCTGCGCCCCACAGCC
CCTTTGGAGCTTTTCTCTTGTCCTCTCACTCCTTCCCAGAAGTTTTTGCACAGAACTTCATTTTGAAAGT
GTTTTTCTCATTCTCCATACCTCCCCCAAGCTCTCCTCCAGCCCTTCCCAGGGCTCAGCCCTGCTGTCCT
GAGCGTCTCCTGGGCCAGAGAGAGGAGATGGGGGTGGAGGGACTGAGTTGATGTTGGGTTTTTCATTCA
ATAAATTGGTGATTTCTTACCGAC

Human PACS-1 protein sequence - var1 (public gi: 7022111) (SEQ ID NO: 362)
MPRPVKVAAVGGQSYLSSILRFFVKSLANMTSDWLGYMRFLIIPLGSHPVAKYLGSVDSKYSSSFLDSGW
RDLFSRSEPPVSEQLDVAGRVMQYVNGAATTHQLPVAEAMLTCRHKFPDEDSYQKFIPFIGVVKVGLVED
SPSTAGDGDDSPVVSLTVPSTSPPSSSGLSRDATATPPSSPSMSSALAIVGSPNSPYGDVIGLQVDYWLG
HPGERRREGDKRDASSKNTLKSVFRSVQVSRLPHSGEAQLSGTMAMTVVTKEKNKKVPTIFLSKKPREKE
VDSKSQVIEGISRLICSSPSLGPSLGPDPSSQPGFPPAGSFPPCHLPLTNPGSEPLIPDRPCSQEWLRTQ
GPSPALCTPQPGHLRPTAPLELFSCPLTPSQKFLHRTSF

Human PACS-1 protein sequence - var2 (public gi: 6330231) (SEQ ID NO: 363)

AITRQPNIKQKFVALLKRFKVSDEVGFGLEHVSREQIREVEEDLDELYDSLEMYNPSDSGPEMEETESIL
STPKPKLKPFFEGMSQSSSQTEIGSLNSKGSLGKDTTSPMELAALEKIKSTWIKNQDDSLTETDTLEITD
QDMFGDASTSLVVPEKVKTPMKSSKTDLQGSASPSKVEGVHTPRQKRSTPLKERQLSKPLSERTNSSDSE
RSPDLGHSTQIPRKVVYDQLNQILVSDAALPENVILVNTTDWQGQYVAELLQDQRKPVVCTCSTVEVQAV
LSALLTRIQRYCNCNSSMPRPVKVAAVGGQSYLSSILRFFVKSLANKTSDWLGYMRFLIIPLGSHPVAKY
LGSVDSKYSSSFLDSGWRDLFSRSEPPVSEQLDVAGRVMQYVNGAATTHQLPVAEAMLTCRHKFPDEDSY
QKFIPFIGVVKVGLVEDSPSTAGDGDDSPVVSLTVPSTSPPSSSGLSRDATATPPSSPSMSSALAIVGSP
NSPYGDVIGLQVDYWLGHPGERRREGDKRDASSKNTLKSVFRSVQVSRLPHSGEAQLSGTMAMTVVTKEK
NKKVPTIFLSKKPREKEVDSKSQVIEGISRLICSAKQQQTMLRVSIDGVEWSDIKFFQLAAQWPTHVKHF
PVGLFSGSKAT

Human PACS-1 protein sequence - var3 (public gi: 34420885) (SEQ ID NO: 364)
MAERGGAGGGPGGAGGGSGQRGSGVAQSPQQPPPQQQQQQPPQQPTPPKLAQATSSSSSTSAAAASSSSS
STSTSMAVAVASGSAPPGGPGPGRTPAPVQMNLYATWEVDRSSSSCVPRLFSLTLKKLVMLKEMDKDLNS
VVIAVKLQGSKRILRSNEIVLPASGLVETELQLTFSLQYPHFLKRDANKLQIMLQRRKRYKNRTILGYKT
LAVGLINMAEVMQHPNEGALVLGLHSNVKDVSVPVAEIKIYSLSSQPIDHEGIKSKLSDRSPDIDNYSEE
EEESFSSEQEGSDDPLHGQDLFYEDEDLRKVKKTRRKLTSTSAITRQPNIKQKFVALLKRFKVSDEVGFG
LEHVSREQIREVEEDLDELYDSLEMYNPSDSGPEMEETESILSTPKPKLKPFFEGMSQSSSQTEIGSLNS
KGSLGKDTTSPMELAALEKIKSTWIKNQDDSLTETDTLEITDQDMFGDASTSLVVPEKVKTPMKSSKTDL
QGSASPSKVEGVHTPRQKRSTPLKERQLSKPLSERTNSSDSERSPDLGHSTQIPRKVVYDQLNQILVSDA
ALPENVILVNTTDWQGQYVAELLQDQRKPVVCTCSTVEVQAVLSALLTRIQRYCNCNSSMPRPVKVAAVG
GQSYLSSILRFFVKSLANKTSDWLGYMRFLIIPLGSHPVAKYLGSVDSKYSSSFLDSGWRDLFSRSEPPV
SEQLDVAGRVMQYVNGAATTHQLPVAEAMLTCRHKFPDEDSYQKFIPFIGVVKVGLVEDSPSTAGDGDDS
PVVSLTVPSTSPPSSSGLSRDATATPPSSPSMNSALAIVGSPNSPYGDVIGLQVDYWLGHPGERRREGDK
RDASSKNTLKSVFRSVQVSRLPHSGEAQLSGTMAMTVVTKELNKKVPTIFLSKKPREKEVDSKSQVIEGI
SRLICSAKQQQTMLRVSIDGVEWSDIKFFQLAAQWPTHVKHFPVGLFSGSKAT

Human PACS-1 protein sequence - var4 (public gi: 33243995) (SEQ ID NO: 365)
ESILSTPKPKLKPFFEGMSQSSSQTEIGSLNSKGSLGKDTTSPMELAALEKIKSTWIKNQDDSLTETDTL
EITDQDMFGDASTSLVVPEKVKTPMKSSKTDLQGSASPSKVEGVHTPRQKRSTPLKERQLSKPLSERTNS
SDSERSPDLGHSTQIPRKVVYDQLNQILVSDAALPENVILVNTTDWQGQYVAELLQDQRKPVVCTCSTVE
VQAVLSALLTRIQRYCNCNSSMPRPVKVAAVGGQSYLSSILRFFVKSLANKTSDWLGYMRFLIIPLGSHP
VAKYLGSVDSKYSSSFLDSGWRDLFSRSEPPVSEQLDVAGRVMQYVNGAATTHQLPVAEAMLTCRHKFPD
EDSYQKFIPFIGVVKVGLVEDSPSTAGDGDDSPVVSLTVPSTSPPSSSGLSRDATATPPSSPSMSSALAI
VGSPNSPYGDVIGLQVDYWLGHPGERRREGDKRDASSKNTLKSVFRSVQVSRLPHSGEAQLSGTMAMTVV
TKEKNKKVPTIFLSKKPREKEVDSKSQVIEGISRLICSAKQQQTMLRVSIDGVEWSDIKFFQLAAQWPTH
VKHFPVGLFSGSKAT

Human PACS-1 protein sequence - var5 (public gi: 30962846) (SEQ ID NO: 366)
MAERGGAGGGPGGAGGGSGQRGSGVAQSPQQPPPQQQQQQPPQQPTPPKLAQATSSSSSTSAAAASSSSS
STSTSMAVAVASGSAPPGGPGPGRTPAPVQMNLYATWEVDRSSSSCVPRLFSLTLKKLVMLKEMDKDLNS
VVIAVKLQGSKRILRSNEIVLPASGLVETELQLTFSLQYPHFLKRDANKLQIMLQRRKRYKNRTILGYKT
LAVGLINMAEVMQHPNEGALVLGLHSNVKDVSVPVAEIKIYSLSSQPIDHEGIKSKLSDRSPDIDNYSEE
EEESFSSEQEGSDDPLHGQDLFYEDEDLRKVKKTRRKLTSTSAITRQPNIKQKFVALLKRFKVSDEVGFG
LEHVSREQIREVEEDLDELYDSLEMYNPSDSGPEMEETESILSTPKPKLKPFFEGMSQSSSQTEIGSLNS
KGSLGKDTTSPMELAALEKIKSTWIKNQDDSLTETDTLEITDQDMFGDASTSLVVPEKVKTPMKSSKTDL
QGSASPSKVEGVHTPRQKRSTPLKERQLSKPLSERTNSSDSERSPDLGHSTQIPRKVVYDQLNQILVSDA
ALPENVILVNTTDWQGQYVAELLQDQRKPVVCTCSTVEVQAVLSALLTRIQRYCNCNSSMPRPVKVAAVG
GQSYLSSILRFFVKSLANKTSDWLGYMRFLIIPLGSHPVAKYLGSVDSKYSSSFLDSGWRDLFSRSEPPV
SEQLDVAGRVMQYVNGAATTHQLPVAEAMLTCRHKFPDEDSYQKFIPFIGVVKVGLVEDSPSTAGDGDDS
PVVSLTVPSTSPPSSSGLSRDATATPPSSPSMSSALAIVGSPNSPYGDVIGLQVDYWLGHPGERRREGDK
RDASSKNTLKSVFRSVQVSRLPHSGEAQLSGTMAMTVVTKEKNKKVPTIFLSKKPREKEVDSKSQVIEGI
SRLICSAKQQQTMLRVSIDGVEWSDIKFFQLAAQWPTHVKHFPVGLFSGSKAT

Unigene Name: PPP1CA Unigene ID: Hs.183994

Human PPP1CA mRNA sequence - var1 (public gi: 287796) (SEQ ID NO: 101)
GCAAGGAGCTGCTGGCTGGACGGCGGCATGTCCGACAGCGAGAAGCTCAACCTGGACTCGATCATCGGGC
GCCTGCTGGAAGTGCAGGGCTCGCGGCCTGGCAAGAATGTACAGCTGACAGAGAACGAGATCCGCGGTCT
GTGCCTGAAATCCCGGGAGATTTTTCTGAGCCAGCCCATTCTTCTGGAGCTGGAGGCACCCCTCAAGATC
TGCGGTGACATACACGGCCAGTACTACGACCTTCTGCGACTATTTGAGTATGGCGGTTTCCCTCCCGAGA

GCAACTACCTCTTTCTGGGGGACTATGTGGACAGGGGCAAGCAGTCCTTGGAGACCATCTGCCTGCTGCT
GGCCTATAAGATCAAGTACCCCGAGAACTTCTTCCTGCTCCGTGGGAACCACGAGTGTGCCAGCATCAAC
CGCATCTATGGTTTCTACGATGAGTGCAAGAGACGCTACAACATCAAACTGTGGAAAACCTTCACTGACT
GCTTCAACTGCCTGCCCATCGCGGCCATAGTGGACGAAAAGATCTTCTGCTGCCACGGAGGCCTGTCCCC
GGACCTGCAGTCTATGGAGCAGATTCGGCGGATCATGCGGCCCACAGATGTGCCTGACCAGGGCCTGCTG
TGTGACCTGCTGTGGTCTGACCCTGACAAGGACGTGCAGGGCTGGGGCGAGAACGACCGTGGCGTCTCTT
TTACCTTTGGAGCCGAGGTGGTGGCCAAGTTCCTCCACAAGCACGACTTGGACCTCATCTGCCGAGCACA
CCAGGTGGTAGAAGACGGCTACGAGTTCTTTGCCAAGCGGCAGCTGGTGACACTTTTCTCAGCTCCCAAC
TACTGTGGCGAGTTTGACAATGCTGGCGCCATGATGAGTGTGGACGAGACCCTCATGTGCTCTTTCCAGA
TCCTCAAGCCCGCCGACAAGAACAAGGGGAAGTACGGGCAGTTCAGTGGCCTGAACCCTGGAGGCCGACC
CATCACCCCACCCCGCAATTCCGCCAAAGCCAAGAAATAGCCCCCGCACACCACCCTGTGCCCCAGATGA
TGGATTGATTGTACAGAAATCATGCTGCCATGCTGGGGGGGGGTCACCCCGACCCCTAAGGCCCACCTGT
CACGGGGAACATGGAGCCTTGGTGTATTTTTCTTTTCTTTTTTTAATGAATCAATAGCAGCGTCCAGTCC
CCCAGGGCTGCTTCCTGCCTGCACCTGCGGTACTGTGAGCAGGATCCTGGGGCCGAGGCTGCAGCTCAGG
GCAACGGCAGGCCAGGTCGTGGGTCTCCAGCCGTGCTTGGCCTCAGGCTGGCAGCCCGGATCCTGGGGCA
ACCCATCTGGTCTCTTGAATAAAGGTCAAAGCTGGATCGGAATC

Human PPP1CA mRNA sequence - var2 (public gi: 21758300) (SEQ ID NO: 102)
AAAAAAAAAAAAAGTTTTCCCTCCATGAGGCAGCGCGCCGACCGCCGAAGCATGGTCTCCACCAGCGGCG
CCGCCACCTCCAGCGTCCTCGGCAGGGAGTTGTGGTGCCGTAGAGGGCGGTCCCCGCGGCCCACGCCGCA
CACCACCTGGGCAGGGGGAGACTCAGGGGGAGGCCCACACACTCCCTGCCCCCCACGCACACCCCTACCG
CCTTGTGCCAAAATTCAGACCAGACCCCTCACTGGACATTCAAGAAGCCCCGTCCTCCAACGTGTCTTAA
ATTGCACACGAGCTCTCCCTGCCACTCCCCATCTGGTCCCCAGACCTCTCCAGGGATTCTACCTACCCAG
GCTTCCAGGCCCAGCTGGGGTCCCCCTCCAGGATGGCTCCTGCAGCCCTGGGGGCCTGGGCCACCCTGGT
GTGCCCCACCCTAGCATCTCCCTGGGGCGCACCTTTCCCTACCCCACTGGAGCTCCCTGAGGGCAGGGTC
GAATCTCTCCCTCTCAGTGTAGCCTAGAGCGGGGTACTCAGGAGGGTCCGTAAGCCTTCCTGACTCTCCA
GCTTAGAGGCCCCTCCTGAAGGCGTCCAGGCACTAGAGGTTTATCAGGAGGCCCTGGGTCAGCCTCTACG
TGGGCAAGAGCTCTCTGGGAAGACGGGGAGGTCTAAGGCCAGCACAGAGTGGCCAGAGGGCCACACCAAA
CTCCCATCCCTGGTCAGCCCAGGTGGCTCTCACCTGAGCAGGGCAGCTGGGCAGGTGGGTACACAGCCTC
CACCAGGACACTCTCCTCCTCCTCCAGCTTCTCCAGCAGCGCCAGCACTGTGTCCACCACTGCACCCAGC
TCTGCCCGCGGGTGCAGACGCCATGCCTGCCGCCCCCGCCAGCGCCAGCCACTGAGCTTCACAGCTACCT
GCAGCAAGGAGGGGAAAGGGGCCTCCTGGACACCACCCCAGGTACTGCAGGGTGGGCACTTCCGCCACA
GGAGCCGTGCAGGGCTCGCGGCCTGGCAAGAATGTACAGCTGACAGAGAACGAGATCCGCGGTCTGTGCC
TGAAATCCCGGGAGATTTTTCTGAGCCAGCCCATTCTTCTGGAGCTGGAGGCACCCCTCAAGATCTGCGG
TGACATACACGGCCAGTACTACGACCTTCTGCGACTATTTGAGTATGGCGGTTTCCCTCCCGAGAGCAAC
TACCTCTTTCTGGGGGACTATGTGGACAGGGGCAAGCAGTCCTTGGAGACCATCTGCCTGCTGCTGGCCT
ATAAGATCAAGTACCCCGAGAACTTCTTCCTGCTCCGTGGGAACCACGAGTGTGCCAGCATCAACCGCAT
CTATGGTTTCTACGATGAGTGCAAGAGACGCTACAACATCAAACTGTGGAAAACCTTCACTGACTGCTTC
AACTGCCTGCCCATCGCGGCCATAGTGGACGAAAAGATCTTCTGCTGCCACGGAGGCCTGTCCCCGGACC
TGCAGTCTATGGAGCAGATTCGGCGGATCATGCGGCCCACAGATGTGCCTGACCAGGGCCTGCTGTGTGA
CCTGCTGTGGTCTGACCCTGACAAGGACGTGCAGGGCTGGGGCGAGAACGACCGTGGCGTCTCTTTTACC
TTTGGAGCCGAGGTGGTGGCCAAGTTCCTCCACAAGCACGACTTGGACCTCATCTGCCGAGCACACCAGG
TGGTAGAAGACGGCTACGAGTTCTTTGCCAAGCGGCAGCTGGTGACACTTTTCTCAGCTCCCAACTACTG
TGGCGAGTTTGACAATGCTGGCGCCATGATGAGTGTGGACGAGACCCTCATGTGCTCTTTCCAGATCCTC
AAGCCCGCCGACAAGAACAAGGGGAAGTACGGGCAGTTCAGTGGCCTGAACCCTGGAGGCCGACCCATCA
CCCCACCCCGCAATTCCGCCAAAGCCAAGAAATAGCCCCCGCACACCACCCTGTGCCCCAGATGATGGAT
TGATTGTACAGAAATCATGCTGCCATGCTGGGGGGGGGTCACCCCGACCCCTCAGGCCCACCTGTCACGG
GGAACATGGAGCCTTGGTGTATTTTTCTTTTCTTTTTTTAATGAATCAATAGCAGCGTCCAGTCCCCCAG
GGCTGCTTCCTGCCTGCACCTGCGGTGACTGTGAGCAGGATCCTGGGGCCGAGGCTGCAGCTCAGGGCAA
CGGCAGGCCAGGTCGTGGGTCTCCAGCCGTGCTTGGCCTCAGGGCTGGCAGCCGGATCCTGGGGCAACCC
ATCTGGTCTCTTGAATAAAGGTCAAAGCTGGATTCTCGC

Human PPP1CA mRNA sequence - var3 (public gi: 14124967) (SEQ ID NO: 103)
GGCTGCCGGAGGGCGGGAGGCAGGAGCGGGCCAGGAGCTGCTGGGCTGGAGCGGCGGCGCCGCCATGTCC
GACAGCGAGAAGCTCAACCTGGACTCGATCATCGGGCGCCTGCTGGAAGTGCAGGGCTCGCGGCCTGGCA
AGAATGTACAGCTGACAGAGAACGAGATCCGCGGTCTGTGCCTGAAATCCCGGGAGATTTTTCTGAGCCA
GCCCATTCTTCTGGAGCTGGAGGCACCCCTCAAGATCTGCGGTGACATACACGGCCAGTACTACGACCTT
CTGCGACTATTTGAGTATGGCGGTTTCCCTCCCGAGAGCAACTACCTCTTTCTGGGGGACTATGTGGACA
GGGGCAAGCAGTCCTTGGAGACCATCTGCCTGCTGCTGGCCTATAAGATCAAGTACCCCGAGAACTTCTT
CCTGCTCCGTGGGAACCACGAGTGTGCCAGCATCAACCGCATCTATGGTTTCTACGATGAGTGCAAGAGA
CGCTACAACATCAAACTGTGGAAAACCTTCACTGACTGCTTCAACTGCCTGCCCATCGCGGCCATAGTGG
ACGAAAAGATCTTCTGCTGCCACGGAGGCCTGTCCCCGGACCTGCAGTCTATGGAGCAGATTCGGCGGAT

CATGCGGCCCACAGATGTGCCTGACCAGGGCCTGCTGTGTGACCTGCTGTGGTCTGACCCTGACAAGGAC
GTGCAGGGCTGGGGCGAGAACGACCGTGGCGTCTCTTTTACCTTTGGAGCCGAGGTGGTGGCCAAGTTCC
TCCACAAGCACGACTTGGACCTCATCTGCCGAGCACACCAGGTGGTAGAAGACGGCTACGAGTTCTTTGC
CAAGCGGCAGCTGGTGACACTTTTCTCAGCTCCCAACTACTGTGGCGAGTTTGACAATGCTGGCGCCATG
ATGAGTGTGGACGAGACCCTCATGTGCTCTTTCCAGATCCTCAAGCCCGCCGACAAGAACAAGGGGAAGT
ACGGGCAGTTCAGTGGCCTGAACCCTGGAGGCCGACCCATCACCCCACCCCGCAATTCCGCCAAAGCCAA
GAAATAGCCCCCGCACACCACCCTGTGCCCCAGATGATGGATTGATTGTACAGAAATCATGCTGCCATGC
TGGGGGGGGGTCACCCCGACCCCTCAGGCCCACCTGTCACGGGGAACATGGAGCCTTGGTGTATTTTTCT
TTTCTTTTTTAATGAATCAATAGCAGCGTCCAGTCCCCCAGGGCTGCTTCCTGCCTGCACCTGCGGTGA
CTGTGAGCAGGATCCTGGGGCCGAGGCTGCAGCTCAGGGCAACGGCAGGCCAGGTCGTGGGTCTCCAGCC
GTGCTTGGCCTCAGGGCTGGCAGCCGGATCCTGGGGCAACCCATCTGGTCTCTTGAATAAAGGTCAAAGC
TGGATTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human PPP1CA mRNA sequence - var4 (public gi: 33872852) (SEQ ID NO: 104)
CCTCGTGCCGAATTCGGCACGAGGAGCGGGCCAGGAGCTGCTGGGCTGGAGCGGCGGCGCCGCCATGTCC
GACAGCGAGAAGCTCAACCTGGACTCGATCATCGGGCGCCTGCTGGAAGTGCAGGGCTCGCGGCCTGGCA
AGAATGTACAGCTGACAGAGAACGAGATCCGCGGTCTGTGCCTGAAATCCCGGGAGATTTTTCTGAGCCA
GCCCATTCTTCTGGAGCTGGAGGCACCCCTCAAGATCTGCGGTGACATACACGGCCAGTACTACGACCTT
CTGCGACTATTTGAGTATGGCGGTTTCCCTCCCGAGAGCAACTACCTCTTTCTGGGGGACTATGTGGACA
GGGGCAAGCAGTCCTTGGAGACCATCTGCCTGCTGCTGGCCTATAAGATCAAGTACCCCGAGAACTTCTT
CCTGCTCCGTGGGAACCACGAGTGTGCCAGCATCAACCGCATCTATGGTTTCTACGATGAGTGCAAGAGA
CGCTACAACATCAAACTGTGGAAAACCTTCACTGACTGCTTCAACTGCCTGCCCATCGCGGCCATAGTGG
ACGAAAAGATCTTCTGCTGCCACGGAGGCCTGTCCCCGGACCTGCAGTCTATGGAGCAGATTCGGCGGAT
CATGCGGCCCACAGATGTGCCTGACCAGGGCCTGCTGTGTGACCTGCTGTGGTCTGACCCTGACAAGGAC
GTGCAGGGCTGGGGCGAGAACGACCGTGGCGTCTCTTTTACCTTTGGAGCCGAGGTGGTGGCCAAGTTCC
TCCACAAGCACGACTTGGACCTCATCTGCCGAGCACACCAGGTGGTAGAAGACGGCTACGAGTTCTTTGC
CAAGCGGCAGCTGGTGACACTTTTCTCAGCTCCCAACTACTGTGGCGAGTTTGACAATGCTGGCGCCATG
ATGAGTGTGGACGAGACCCTCATGTGCTCTTTCCAGATCCTCAAGCCCGCCGACAAGAACAAGGGGAAGT
ACGGGCAGTTCAGTGGCCTGAACCCTGGAGGCCGACCCATCACCCCACCCCGCAATTCCGCCAAAGCCAA
GAAATAGCCCCCGCACACCACCCTGTGCCCCAGATGATGGATTGATTGTACAGAAATCATGCTGCCATGC
TGGGGGGGGGGTCACCCCGACCCCTCAGGCCCACCTGTCACGGGGAACATGGAGCCTTGGTGTATTTTTC
TTTTCTTTTTTTAATGAATCAATAGCAGCGTCCAGTCCCCCAGGGCTGCTTCCTGCCTGCACCTGCGGTG
ACTGTGAGCAGGATCCTGGGGCCGAGGCTGCAGCTCAGGGCAACGGCAGGCCAGGTCGTGGGTCTCCAGC
CGTGCTTGGCCTCAGGGCTGGCAGCCGGATCCTGGGGCAACCCATCTGGTCTCTTGAATAAAGGTCAAAG
CTGGATTCTCGAAAAAAAAAAAAAAAAAA

Human PPP1CA mRNA sequence - var5 (public gi: 12804878) (SEQ ID NO: 105)
CAGGAGCGGGCCAGGAGCTGCTGGGCTGGAGCGGCGGCGCCGCCATGTCCGACAGCGAGAAGCTCAACCT
GGACTCGATCATCGGGCGCCTGCTGGAAGTGCAGGGCTCGCGGCCTGGCAAGAATGTACAGCTGACAGAG
AACGAGATCCGCGGTCTGTGCCTGAAATCCCGGGAGATTTTTCTGAGCCAGCCCATTCTTCTGGAGCTGG
AGGCACCCCTCAAGATCTGCGGTGACATACACGGCCAGTACTACGACCTTCTGCGACTATTTGAGTATGG
CGGTTTCCCTCCCGAGAGCAACTACCTCTTTCTGGGGGACTATGTGGACAGGGGCAAGCAGTCCTTGGAG
ACCATCTGCCTGCTGCTGGCCTATAAGATCAAGTACCCCGAGAACTTCTTCCTGCTCCGTGGGAACCACG
AGTGTGCCAGCATCAACCGCATCTATGGTTTCTACGATGAGTGCAAGAGACGCTACAACATCAAACTGTG
GAAAACCTTCACTGACTGCTTCAACTGCCTGCCCATCGCGGCCATAGTGGACGAAAAGATCTTCTGCTGC
CACGGAGGCCTGTCCCCGGACCTGCAGTCTATGGAGCAGATTCGGCGGATCATGCGGCCCACAGATGTGC
CTGACCAGGGCCTGCTGTGTGACCTGCTGTGGTCTGACCCTGACAAGGACGTGCAGGGCTGGGGCGAGAA
CGACCGTGGCGTCTCTTTTACCTTTGGAGCCGAGGTGGTGGCCAAGTTCCTCCACAAGCACGACTTGGAC
CTCATCTGCCGAGCACACCAGGTGGTAGAAGACGGCTACGAGTTCTTTGCCAAGCGGCAGCTGGTGACAC
TTTTCTCAGCTCCCAACTACTGTGGCGAGTTTGACAATGCTGGCGCCATGATGAGTGTGGACGAGACCCT
CATGTGCTCTTTCCAGATCCTCAAGCCCGCCGACAAGAACAAGGGGAAGTACGGGCAGTTCAGTGGCCTG
AACCCTGGAGGCCGACCCATCACCCCACCCCGCAATTCCGCCAAAGCCAAGAAATAGCCCCCGCACACCA
CCCTGTGCCCCAGATGATGGATTGATTGTACAGAAATCATGCTGCCATGCTGGGGGGGGGTCACCCCGAC
CCCTCAGGCCCACCTGTCACGGGGAACATGGAGCCTTGGTGTATTTTTCTTTTCTTTTTTTAATGAATCA
ATAGCAGCGTCCAGTCCCCCAGGGCTGCTTCCTGCCTGCACCTGCGGTGACTGTGAGCAGGATCCTGGGG
CCGAGGCTGCAGCTCAGGGCAACGGCAGGCCAGGTCGTGGGTCTCCAGCCGTGCTTGGCCTCAGGGCTGG
CAGCCGGATCCTGGGGCAACCCATCTGGTCTCTTGAATAAAGGTCAAAGCTGGATTCTCAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human PPP1CA mRNA sequence - var6 (public gi: 34534606) (SEQ ID NO: 106)
CTTCTTGCTGACGCCGCCAGCGCCGACCACCAGCAGCTGTTTTCCCTCCATGAGGCAGCGCGCCGACCGC
CGAAGCATGGTCTCCACCAGCGGCGCCGCCACCGCCTCGTCGGCCGCCGGCCCCAGCCGCGGCGCGGCCC

ACAGCCCCTCCAGCGCGCCGCACGCCTCCAGACACAGGCCGCCGTTCAGCTCCAGGGCCACTGGGCTTCT
CCAGCAGCGCCAGCACTGTGTCCACCACTGCACCCAGCTCTGCCCGCGGGTGCAGACGCCATGCCTGCCG
CCCCCGCCAGCGCCAGCCACTGAGCTTCACAGCTACCTGCAGCAAGGAGGGGAAAGGGGCCTCCTGGACA
CCACCCCAGGTACTGCAGGGTGGGGCACTTCCGCCACAGGAGCCGTGCAGGGCTCGCGGCCTGGCAAGAA
TGTACAGCTGACAGAGAACGAGATCCGCGGTCTGTGCCTGAAATCCCGGGAGATTTTTCTGAGCCAGCCC
ATTCTTCTGGAGCTGGAGGCACCCCTCAAGATCTGCGGTGACATACACGGCCAGTACTACGACCTTCTGC
GACTATTTGAGTATGGCGGTTTCCCTCCCGAGAGCAACTACCTCTTTCTGGGGGACTATGTGGACAGGGG
CAAGCAGTCCTTGGAGACCATCTGCCTGCTGCTGGCCTATAAGATCAAGTACCCCGAGAACTTCTTCCTG
CTCCGTGGGAACCACGAGTGTGCCAGCATCAACCGCATCTATGGTTTCTACGATGAGTGCAAGAGACGCT
ACAACATCAAACTGTGGAAAACCTTCACTGACTGCTTCAACTGCCTGCCCATCGCGGCCATAGTGGACGA
AAAGATCTTCTGCTGCCACGGAGGCCTGTCCCCGGACCTGCAGTCTATGGAGCAGATTCGGCGGATCATG
CGGCCCACAGATGTGCCTGACCAGGGCCTGCTGTGTGACCTGCTGTGGTCTGACCCTGACAAGGACGTGC
AGGGCTGGGGCGAGAACGACCGTGGCGTCTCTTTTACCTTTGGAGCCGAGGTGGTGGCCAAGTTCCTCCA
CAAGCACGACTTGGACCTCATCTGCCGAGCACACCAGGTGGTAGAAGACGGCTACGAGTTCTTTGCCAAG
CGGCAGCTGGTGGCACTTTTCTCAGCTCCCAACTACTGTGGCGAGTTTGACAATGCTGGCGCCATGATGA
GTGTGGACGAGACCCTCATGTGCTCTTTCCAGATCCTCAAGCCCGCCGACAAGAACAAGGGGAAGTACGG
GCAGTTCAGTGGCCTGAACCCTGGAGGCTGACCCATCACCCCACCCCGCAATTCCGCCAAAGCCAAGAAA
TAGCCCCCGCACACCACCCTGTGCCCCAGATGATGGATTGATTGTACAGAAATCATGCTGCCATGCTGGG
GGGGGGTCACCCCGACCCCTCAGGCCCACCTGTCACGGGGAACATGGAGCCTTGGTGTATTTTTCTTTTC
TTTTTTTAATGAATCAATAGCAGCGTCCAGTCCCCCAGGGCTGCTTCCTGCCTGCACCTGCGGTGACTGT
GAGCAGGATCCTGGGGCCGAGGCTGCAGCTCAGGGCAACGGCAGGCCAGGTCGTGGGTCTCCAGCCGTGC
TTGGCCTCAGGGCTGGCAGCCGGATCCTGGGGCAACCCATCTGGTCTCTTGAATAAAGGTCAAAGCTGGA
TTCTC

Human PPP1CA mRNA sequence - var7 (public gi: 30582096) (SEQ ID NO: 107)
ATGTCCGACAGCGAGAAGCTCAACCTGGACTCGATCATCGGGCGCCTGCTGGAAGTGCAGGGCTCGCGGC
CTGGCAAGAATGTACAGCTGACAGAGAACGAGATCCGCGGTCTGTGCCTGAAATCCCGGGAGATTTTTCT
GAGCCAGCCCATTCTTCTGGAGCTGGAGGCACCCCTCAAGATCTGCGGTGACATACACGGCCAGTACTAC
GACCTTCTGCGACTATTTGAGTATGGCGGTTTCCCTCCCGAGAGCAACTACCTCTTTCTGGGGGACTATG
TGGACAGGGGCAAGCAGTCCTTGGAGACCATCTGCCTGCTGCTGGCCTATAAGATCAAGTACCCCGAGAA
CTTCTTCCTGCTCCGTGGGAACCACGAGTGTGCCAGCATCAACCGCATCTATGGTTTCTACGATGAGTGC
AAGAGACGCTACAACATCAAACTGTGGAAAACCTTCACTGACTGCTTCAACTGCCTGCCCATCGCGGCCA
TAGTGGACGAAAAGATCTTCTGCTGCCACGGAGGCCTGTCCCCGGACCTGCAGTCTATGGAGCAGATTCG
GCGGATCATGCGGCCCACAGATGTGCCTGACCAGGGCCTGCTGTGTGACCTGCTGTGGTCTGACCCTGAC
AAGGACGTGCAGGGCTGGGGCGAGAACGACCGTGGCGTCTCTTTTACCTTTGGAGCCGAGGTGGTGGCCA
AGTTCCTCCACAAGCACGACTTGGACCTCATCTGCCGAGCACACCAGGTGGTAGAAGACGGCTACGAGTT
CTTTGCCAAGCGGCAGCTGGTGACACTTTTCTCAGCTCCCAACTACTGTGGCGAGTTTGACAATGCTGGC
GCCATGATGAGTGTGGACGAGACCCTCATGTGCTCTTTCCAGATCCTCAAGCCCGCCGACAAGAACAAGG
GGAAGTACGGGCAGTTCAGTGGCCTGAACCCTGGAGGCCGACCCATCACCCCACCCCGCAATTCCGCCAA
AGCCAAGAAATAG

Human PPP1CA mRNA sequence - var8 (public gi: 190515) (SEQ ID NO: 108)
GGGCAAGGAGCTGCTGGCTGGACGGCGGCATGTCCGACAGCGAGAAGCTCAACCTGGACTCGATCATCGG
GCGCCTGCTGGAAGTGCAGGGCTCGCGGCCTGGCAAGAATGTACAGCTGACAGAGAACGAGATCCGCGGT
CTGTGCCTGAAATCCCGGGAGATTTTTCTGAGCCAGCCCATTCTTCTGGAGCTGGAGGCACCCCTCAAGA
TCTGCGGTGACATACACGGCCAGTACTACGACCTTCTGCGACTATTTGAGTATGGCGGTTTCCCTCCCGA
GAGCAACTACCTCTTTCTGGGGGACTATGTGGACAGGGGCAAGCAGTCCTTGGAGACCATCTGCCTGCTG
CTGGCCTATAAGATCAAGTACCCCGAGAACTTCTTCCTGCTCCGTGGGAACCACGAGTGTGCCAGCATCA
ACCGCATCTATGGTTTCTACGATGAGTGCAAGAGACGCTACAACATCAAACTGTGGAAAACCTTCACTGA
CTGCTTCAACTGCCTGCCCATCGCGGCCATAGTGGACGAAAAGATCTTCTGCTGCCACGGAGGCCTGTCC
CCGGACCTGCAGTCTATGGAGCAGATTCGGCGGATCATGCGGCCCACAGATGTGCCTGACCAGGGCCTGC
TGTGTGACCTGCTGTGGTCTGACCCTGACAAGGACGTGCAGGGCTGGGGCGAGAACGACCGTGGCGTCTC
TTTTACCTTTGGAGCCGAGGTGGTGGCCAAGTTCCTCCACAAGCACGACTTGGACCTCATCTGCCGAGCA
CACCAGGTGGTAGAAGACGGCTATGAGTTCTTTGCCAAGCGGCAGCTGGTGACACTTTTCTCAGCTCCCA
ACTACTGTGGCGAGTTTGACAATGCTGGCGCCATGATGAGTGTGGACGAGACCCTCATGTGCTCTTTCCA
GATCCTCAAGCCCGCCGACAAGAACAAGGGGAAGTACGGGCAGTTCAGTGGCCTGAACCCTGGAGGCCGA
CCCATCACCCCACCCCGCAATTCCGCCAAAGCCAAGAAATAGCCCCCGCACACCACCCTGTGCCCCAGAT
GATGGATTGATTGTACAGAAATCATGCTGCCATGCTGGGGGGGGGTCACCCCGACCCCTAAGGCCCACCT
GTCACGGGGAACATGGAGCCTTGGTGTATTTTTCTTTTCTTTTTTTAATGAATCAATAGCAGCGTCCAGT
CCCCCAGGGCTGCTTCCTGCCTGCACCTGCGGTACTGTGAGCAGGATCCTGGGGCCGAGGCTGCAGCTCA
GGGCAACGGCAGGCCAGGTCGTGGGTCTCCAGCCGTGCTTGGCCTCAGGGCTGGCAGCCCGGATCCTGGGG
CAACCCATCTGGTCTCTTGAATAAAGGTCAAAGCTGG

Human PPP1CA mRNA sequence - var9 (public gi: 190280) (SEQ ID NO: 109)
CGGCCTGGCAAGAATGTACAGCTGACAGAGAACGAGATCCGCGGTCTGTGCCTGAAATCCCGGGAGATTT
TTCTGAGCCAGCCCATTCTTCTGGAGCTGGAGGCACCCCTCAAGATCTGCGGTGACATACACGGCCAGTA
CTACGACCTTCTGCGACTATTTGAGTATGGAGGTTTCCCTCCCGAGAGCAACTACCTCTTTCTGGGGGAC
TATGTGGACAGGGGCAAGCAGTCCTTGGAGACCATCTGCCTGCTGCTGGCCTATAAGATCAAGTACCCCG
AGAACTTCTTCCTGCTCCGTGGGAACCACGAGTGTGCCAGCATCAACCGCATCTATGGTTTCTACGATGA
GTGCAAGAGACGCTACAACATCAAACTGTGGAAAACCTTCACTGACTGCTTCAACTGCCTGCCCATCGCG
GCCATAGTGGACGAAAAGATCTTCTGCTGCCACGGAGGCCTGTCCCCGGACCTGCAGTCTATGGAGCAGA
TTCGGCGGATCATGCGGCCCACAGATGTGCCTGACCAGGGCCTGCTGTGTGACCTGCTGTGGTCTGACCC
TGACAAGGACGTGCAGGGCTGGGGCGAGAACGACCGTGGCGTCTCTTTTACCTTTGGAGCCGAGGTGGTG
GCCAAGTTCCTCCACAAGCACGACTTGGACCTCATCTGCCGAGCACACCAGGTGGTAGAAGACGGCTACG
AGTTCTTTGCCAAGCGGCAGCTGGTGACACTTTTCTCAGCTCCCAACTACTGTGGCGAGTTTGACAATGC
TGGCGCCATGATGAGTGTGGACGAGACCCTCATGTGCTCTTTCCAGATCCTCAAGCCCGCCGACAAGAAC
AAGGGGAAGTACGGGCAGTTCAGTGGCCTGAACCCTGGAGGCCGACCCATCACCCCACCCCGCAATTCCG
CCAAAGCCAAGAAATAGCCCCCGCACACCACCCTGTGCCCCAGATGATGGATTGATTGTACAGAAATCAT
GCTGCCATGCTGGGGGGGGGTCACCCCGACCCCTCAGGCCCACCTGTCACGGGGAACATGGACCTTGGTG
TATTTTTCTTTTCTTTTTTTAATGAATCAG

Human PPP1CA protein sequence - var1 (public gi: 298964) (SEQ ID NO: 261)
MSDSEKLNLDSIIGRLLEGSRVLTPHCAPVQGSRPGKNVQLTENEIRGLCLKSREIFLSQPILLELEAPL
KICGDIHGQYYDLLRLFEYGGFPPESNYLFLGDYVDRGKQSLETICLLLAYKIKYPENFFLLRGNHECAS
INRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQSMEQIRRIMRPTDVPDQG
LLCDLLWSDPDKDVQGWGENDRGVSFTFGAEVVAKFLHKHDLDLICRAHQVVEDGYEFFAKRQLVTLFSA
PNYCGEFDNAGAMMSVDETLMCSFQILKPADKNKGKYGQFSGLNPGGRPITPPRNSAKAKK

Human PPP1CA protein sequence - var2 (public gi: 190516) (SEQ ID NO: 262)
MSDSEKLNLDSIIGRLLEVQGSRPGKNVQLTENEIRGLCLKSREIFLSQPILLELEAPLKICGDIHGQYY
DLLRLFEYGGFPPESNYLFLGDYVDRGKQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDEC
KRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQSMEQIRRIMRPTDVPDQGLLCDLLWSDPD
KDVQGWGENDRGVSFTFGAEVVAKFLHKHDLDLICRAHQVVEDGYEFFAKRQLVTLFSAPNYCGEFDNAG
AMMSVDETLMCSFQILKPADKNKGKYGQFSGLNPGGRPITPPRNSAKAKK

Human PPP1CA protein sequence - var3 (public gi: 190281) (SEQ ID NO: 263)
RPGKNVQLTENEIRGLCLKSREIFLSQPILLELEAPLKICGDIHGQYYDLLRLFEYGGFPPESNYLFLGD
YVDRGKQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIA
AIVDEKIFCCHGGLSPDLQSMEQIRRIMRPTDVPDQGLLCDLLWSDPDKDVQGWGENDRGVSFTFGAEVV
AKFLHKHDLDLICRAHQVVEDGYEFFAKRQLVTLFSAPNYCGEFDNAGAMMSVDETLMCSFQILKPADKN
KGKYGQFSGLNPGGRPITPPRNSAKAKK

Human PPP1CA protein sequence - (public gi: 35451) (SEQ ID NO: 395)
MSDSEKLNLDSIIGRLLEVQGSRPGKNVQLTENEIRGLCLKSREIFLSQPILLELEAPLKICGDIHGQYY
DLLRLFEYGGFPPESNYLFLGDYVDRGKQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDEC
KRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQSMEQIRRIMRPTDVPDQGLLCDLLWSDPD
KDVQGWGENDRGVSFTFGAEVVAKFLHKHDLDLICRAHQVVEDGYEFFAKRQLVTLFSAPNYCGEFDNAG
AMMSVDETLMCSFQILKPADKNKGKYGQFSGLNPGGRPITPPRNSAKAKK

Human PPP1CA pray sequence - var1 (SEQ ID NO: 110)
CCGCCTGGTNCTACCCATGACNCACNTACCANTATTACGTCTACATATGGCTCATGGCAGGCCAGTTGAA
ATTCCACACACAATACAAGTGCGCTCATCGACACGCAGAAGAAGGNCATTTTGNTTGNGNAACTTNATTA
TAGGGCNAGNGCCCCCNTGGANTTCCNNTACAACNTNCCAGGATNACGCTCATATGGCCATGGAGGCCAG
TGAATTCCACCCAAGCGGTGGTATCAACGCACAGTGGCCATTATGGCGGGCAGTGGCCANAACCCTGGAG
GCCGACCCATCACCCCACCCCGCAATTCCGCCAAAGCCAAGAAATAGNNGGCGCACACCACCCTGTGCCT
TNNATGATGGATTGATTGTACAGAAATCATGCTGCCATGCTGGGGGGGG

Unigene Name: PRKAR1A **Unigene ID:** Hs.280342

Human PRKAR1A mRNA sequence - var1 (public gi: 34530409) (SEQ ID NO: 111)
ATCGCAGAGTGGAGCGGGGCTGGGAGCAAAGCGCTGAGGGAGCTCGGTACGCCGCCGCCTCGCACCCGCA

GCCTCGCGCCCGCCGCCGCCCGTCCCCAGAGAACCATGGAGTCTGGCAGTACCGCCGCCAGTGAGGAGGC
ACGCAGCCTTCGAGAATGTGAGCTCTACGTCCAGAAGCATAACATTCAAGCGCTGCTCAAAGATTCTATT
GTGCAGTTGTGCACTGCTCGACCTGAGAGACCCATGGCATTCCTCAGGGAATACTTTGAGAGGAGGAGGC
AAAACAGATTCAGAATCTGCAGAAAGCAGGCACTCGTACAGACTCAAGGGAGGATGAGATTTCTCCTCCT
CCACCCAACCCAGTGGTTAAAGGTAGGAGGCGACGAGGTGCTATCAGCGCTGAGGTCTACACGGAGGAAG
ATGCGGCATCCTATGTTAGAAAGGTTATACCAAAAGATTACAAGACAATGGCCGCTTTAGCCAAAGCCAT
TGAAAAGAATGTGCTGTTTTCACATCTTGATGATAATGAGAGAAGTGATATTTTTGATGCCATGTTTTCG
GTCTCCTTTATCGCAGGAGAGACTGTGATTCAGCAAGGTGATGAAGGGGATAACTTCTATGTGATTGATC
AAGGAGAGACGGATGTCTATGTTAACAATGAATGGGCAACCAGTGTTGGGGAAGGAGGGAGCTTTGGAGA
ACTTGCTTTGATTTATGGAACACCGAGAGCAGCCACTGTCAAAGCAAAGACAAATGTGAAATTGTGGGGC
ATCGACCGAGACAGCTATAGAAGAATCCTCATGGGAAGCACACTGAGAAAGCGGAAGATGTATGAGGAAT
TCCTTAGTAAAGTCTCTATTTTAGAGTCTCTGGACAAGTGGGAACGTCTTACGGTAGCTGATGCATTGGA
ACCAGTGCAGTTTGAAGATGGGCAGAAGATTGTGGTGCAGGGAGAACCAGGGGATGAGTTCTTCATTATT
TTAGAGGGGTCAGCTGCTGTGCTACAACGTCGGTCAGAAAATGAAGAGTTTGTTGAAGTGGGAAGATTGG
GGCCTTCTGATTATTTTGGTGAAATTGCACTACTGATGAATCGTCCTCGTGCTGCCACAGTTGTTGCTCG
TGGCCCCTTGAAGTGCGTTAAGCTGGACCGACCTAGATTTGAACGTGTTCTTGGCCCATGCTCAGACATC
CTCAAACGAAACATCCAGCAGTACAACAGTTTTGTGTCACTGTCTGTCTGAAATCTGCCTCCTGTGCCTC
CCTTTTCTCCTCTCCCCAATCCATGCTTCACTCATGCAAACTGCTTTATTTTCCCTACTTGCAGCGCCAA
GTGGCCACTGGCATCGCAGCTTCCTGTCTGTTTATATATTGAAAGTTGCTTTTATTGCACCATTTTCAAT
TTGGAGCATTAACTAAATGCTCATACACAGTTAAATAAATAGAAAGAGTTCTATGGAGACTTTGCTGTTA
CTGCTTCTCTTTGTGCAGTGTTAGTATTCACCCTGGGCAGTGAGTGCCATGCTTTTTGGTGAGGGCAGAT
CCCAGCACCTATTGAATTACCATAGAGTAATGATGTAACAGTGCAAGATTTTTTTTTTTAAGTGACATAA
TTGTCCAGTTATAAGCGTATTTAGACTGTGGCCATATATGCTGTATTTCTTTGTAGAATAAATGGTTTCT
CATTAAACTCTAAAGATTAGGGAAAATGGATATAGAAAATCTTAGTATAGTAGAAAGACATCTGCCTGTA
ATTAAACTAGTTTAAGGGTGGAAAAATGCCCATTTTTGCTAATTATCAATGGGATATGATTGGTTCAGTT
TTTTTTTTTTTCCAGAGTTGTTGTTTGCCAAGCTAATCTGCCTGGTTTTATTTATATCTTGTTATTAATG
TTTCTTCTCCAATTCTGAAATACTTTTGAGTATGGCTATCTATACCTGCCTTTTAAGTTTGAAACTAACT
CATAGATTGCAAATATTGGTTAGTATTTAACTACATCTGCCTCGGCTCACAAATTCCGATTAGACCTTTA
TCCAGCTAGTGCCAAATAATTGATCAGATGCTGAATTGAGAATAAGAATTTGAGGTCTACATTCTTGGTT
GTTAATTTAGAGCGTTTGGTTAAAGTATGTCCTTCAGCTGACTCCAGTATAATCTCCTCTGCTCATTAAA
CTGATTCCAGGAGATTGGATTTGCTGTGACTAGATACAGATGGAGCAAATGTCCTAACAGAGAAATAGAG
GTGATGCTGCTAAAGGGAGAAATGCCAGGCGGACAAAGTTCAGTGTCGGGAATTTTCCCCGTGACATTCA
CTGGGGCATGAGATTTTGGAAGAAGTTTTTACTTTGGTTTAGTCTTTTTTTCCTTCCTTTTTATTCAGC
TAGAATTTCTGGTGGGTTGATGGTAGGGTATAATGTGTCTGTGTTGCTTCAAATTGGTCTGAAAGGCTAT
CCTGCTGAAAGTCCTGCTTTCCTATCTAGCATTTATTCCTCTGGCAAACTTTTCTTTCTTTTCTTTTTTA
AAGTAAACTTGTGTATTGAGTCTTAACTGTATTTCAGTATTTTCCAGCCTTATGTGTTACATTATTCCAA
TGATACCCAACAGTTTATTTTTATTATTTTTTTAAACAAAATTTCACAGTTCTGTAATGTAGGCACTTTT
ATTTTCATTGTGATTTATATATAAGGTAATGTAGGGTTATATTTGGGAGTGACTGCAAGCATTTTTCCAT
CTGTGTGCAACTAACTGACTCTGTTATTGATCCCTTCTCCTGCCCTTTCCCAGGTAATTTAAATTGGTCA
TGGTAGATTTTTTTCATAGATTTGAAAAACTTTTAGGTTGTTACCAAGTATGAAGTATAAATCTGGGGAA
GAGGTTTTATTTACATTTTAGGGTGGGTAAGAAAGCCACCTTGTTACAAATTTTTTAATTTCCAAAATAA
TCTATATTAAATGAGGGTTTCTGATCTGTACTTTGTGTTTAGCTACCTTTTTATATTTAAAAAAATTAAAA
ATGAAAATTATGTTCTTACAAGCTTAAAGCTTGATTTGATCTTTGTTTAAATGCCAAAATGTACTTAAAT
GAGTTACTTAGAATGCCATAAAATTGCAGTTTCATGTATGTATATAATCATGCTCATGTATATTTAGTTA
CGTATAATGCTTTCTGAGTGAGTTTTACTCTTAAATCATTTGGTTAAATCATTTGGCTTGCTGTTTACTC
CCTTCTGTAGTTTTTAATTAAAAGCTTTAAAGATAAGTCTACATTAAACAATGATCACATCTAAAGCTTT
ATCTTTGTGTAATCTAAGTATATGTGAGAAATCAGAATTGGCATAATTTGTCTTAGTTGATATTCAAGGC
TTTAAAAGTCATTATTCCTGGGCTTGGTAAGTGAATTTATGAGATTTACTGCTCTAGAAAGTATAGATGG
CCAAAGGACCGTTTTGTATTGCTTCCTGATTACCAGTCTGATTATACCATGTGTGCTAATATACTTTTTT
TGTTATAGATTGTCTTAATGGTAGGTCAAGTAATAAAAAGAGATGAAATAATTT

Human PRKAR1A mRNA sequence - var2 (public gi: 4884279) (SEQ ID NO: 112)
TATTTTCCAGCCTTATGTGTTACATTATTCCAATGATACCCAACAGTTTATTTTTATTATTTTTTTAAAC
AAAATTTCACAGTTCTGTAATGTAGGCACTTTTATTTTCATTGTGATTTATATATAAGGTAATGTAGGGT
TATATTTGGGAGTGACTGCAAGCATTTTTCCATCTGTGTGCAACTAACTGACTCTGTTATTGATCCCTTC
TCCTGCCCTTTCCCAGGTAATTTAAATTGGTCATGGTAGATTTTTTTCATAGATTTGAAAAACTTTTAGG
TTGTTACCAAGTATGAAGTATAAATCTGGGGAAGAGGTTTTATTTACATTTTAGGGTGGGTAAGAAAGCC
ACCTTGTTACAAATTTTTTAATTTCCAAAATAATCTATATTAAATGAGGGTTTCTGATCTGTACTTTGTG
TTTAGCTACCTTTTTATATTTAAAAAAATTAAAAATGAAAATTACGTTCTTACAAGCTTAAAGCTTGATTT
GATCTTTGTTTAAATGCCAAAATGTACTTAAATGAGTTACTTAGAATGCCATAAAATTGCAGTTTCATGT
ATGTATATAATCATGCTCATGTATATTTAGTTACGTATAATGCTTTCTGAGTGAGTTTTACTCTTAAATC
ATTTGGTTAAATCATTTGGCTTGCTGTTTACTCCCTTCTGTAGTTTTTAATTAAAAACTTTAAAGATAAG
TCTACATTAAACAATGATCACATCTAAAGCTTTATCTTTGTGTAATCTAAGTATATGTGAGAAATCAGAA

TTGGCATAATTTGTCTTAGTTGATATTCAAGGCTTTAAAAGTCATTATTCCTGGGCTTGGTAAGTGAATT
TATGAGATTTACTGCTCTAGAAAGTATAGATGGCCAAAGGACCGTTATGTATTGCTTCCTGATTACCAGT
CTGATTATACCATGTGTGCTAATATACTTTTTTTGTTATAGATTGTCTTAATGGTAGGTCAAGTAATAAA
AAGAGATGAAATAATTTAAAAAAAAAAAAA

Human PRKAR1A mRNA sequence - var3 (public gi: 33636720) (SEQ ID NO: 113)
GGTGGAGCTGTCGCCTAGCCGCTATCGCAGAGTGGAGCGGGGCTGGGAGCAAAGCGCTGAGGGAGCTCGG
TACGCCGCCGCCTCGCACCCGCAGCCTCGCGCCCGCCGCCGCCCGTCCCCAGAGAACCATGGAGTCTGGC
AGTACCGCCGCCAGTGAGGAGGCACGCAGCCTTCGAGAATGTGAGCTCTACGTCCAGAAGCATAACATTC
AAGCGCTGCTCAAAGATTCTATTGTGCAGTTGTGCACTGCTCGACCTGAGAGACCCATGGCATTCCTCAG
GGAATACTTTGAGAGGTTGGAGAAGGAGGAGGCAAAACAGATTCAGAATCTGCAGAAAGCAGGCACTCGT
ACAGACTCAAGGGAGGATGAGATTTCTCCTCCTCCACCCAACCCAGTGGTTAAAGGTAGGAGGCGACGAG
GTGCTATCAGCGCTGAGGTCTACACGGAGGAAGATGCGGCATCCTATGTTAGAAAGGTTATACCAAAAGA
TTACAAGACAATGGCCGCTTTAGCCAAAGCCATTGAAAAGAATGTGCTGTTTTCACATCTTGATGATAAT
GAGAGAAGTGATATTTTTGATGCCATGTTTTCGGTCTCCTTTATCGCAGGAGAGACTGTGATTCAGCAAG
GTGATGAAGGGGATAACTTCTATGTGATTGATCAAGGAGAGACGGATGTCTATGTTAACAATGAATGGGC
AACCAGTGTTGGGGAAGGAGGGAGCTTTGGAGAACTTGCTTTGATTTATGGAACACCGAGAGCAGCCACT
GTCAAAGCAAAGACAAATGTGAAATTGTGGGGCATCGACCGAGACAGCTATAGAAGAATCCTCATGGGAA
GCACACTGAGAAAGCGGAAGATGTATGAGGAATTCCTTAGTAAAGTCTCTATTTTAGAGTCTCTGGACAA
GTGGGAACGTCTTACGGTAGCTGATGCATTGGAACCAGTGCAGTTTGAAGATGGGCAGAAGATTGTGGTG
CAGGGAGAACCAGGGGATGAGTTCTTCATTATTTTAGAGGGGTCAGCTGCTGTGCTACAACGTCGGTCAG
AAAATGAAGAGTTTGTTGAAGTGGGAAGATTGGGGCCTTCTGATTATTTTGGTGAAATTGCACTACTGAT
GAATCGTCCTCGTGCTGCCACAGTTGTTGCTCGTGGCCCCTTGAAGTGCGTTAAGCTGGACCGACCTAGA
TTTGAACGTGTTCTTGGCCCATGCTGAGACATCCTCAAACGAAACATCCAGCAGTACAACAGTTTTGTGT
CACTGTCTGTCTGAAATCTGCCTCCTGTGCCTCCCTTTTCTCCTCTCCCCAATCCATGCTTCACTCATGC
AAACTGCTTTATTTCCCTACTTGCAGCGCCAAGTGGCCACTGGCATCGCAGCTTCCTGTCTGTTTATAT
ATTGAAAGTTGCTTTTATTGCACCATTTTCAATTTGGAGCATTAACTAAATGCTCATACACAGTTAAATA
AATAGAAAGAGTTCTATGGAGACTTTGCTGTTACTGCTTCTCTTTGTGCAGTGTTAGTATTCACCCTGGG
CAGTGAGTGCCATGCTTTTTGGTGAGGGCAGATCCCAGCACCTATTGAATTACCATAGAGTAATGATGTA
ACAGTGCAAGATTTTTTTTTTAAGTGACATAATTGTCCAGTTATAAGCGTATTTAGACTGTGGCCATATA
TGCTGTATTTCTTTGTAGAATAAATGGTTTCTCATTAAACTCTAAAGATTAGGGAAAATGGATATAGAAA
ATCTTAGTATAGTAGAAAGACATCTGCCTGTAATTAAACTAGTTTAAGGGTGGAAAAATGCCCATTTTTG
CTAATTATCAATGGGATATGATTGGTTCAGTTTTTTTTTTTCCAGAGTTGTTGTTTGCCAAGCTAATCTG
CCTGGTTTTATTTATATCTTGTTATTAATGTTTCTTCTCCAATTCTGAAATACTTTTGAGTATGGCTATC
TATACCTGCCTTTTAAGTTTGAAACTAACTCATAGATTGCAAATATTGGTTAGTATTTAACTACATCTGC
CTCGGCTCACAAATTCCGATTAGACCTTTATCCAGCTAGTGCCAAATAATTGATCAGATGCTGAATTGAG
AATAAGAATTTGAGGTCTACATTCTTGGTTGTTAATTTAGAGCGTTTGGTTAAAGTATGTCCTTCAGCTG
ACTCCAGTATAATCTCCTCTGCTCATTAAACTGATTCCAGGAGATTGGATTTGCTGTGACTAGATACAGA
TGGAGCAAATGTCCTAACAGAGAAATAGAGGTGATGCTGCTAAAGGGAGAAATGCCAGGCGGACAAAGTT
CAGTGTCGGGAATTTTCCCCGTGACATTCACTGGGGCATGAGATTTTGGAAGAAGTTTTTTACTTTGGTT
TAGTCTTTTTTTCCTTCCTTTTATTCAGCTAGAATTTCTGGTGGGTTGATGGTAGGGTATAATGTGTCT
GTGTTGCTTCAAATTGGTCTGAAAGGCTATCCTGCGGAAAGTCCTGCTTTCCTATCTAGCATTTATTTCT
CTGGCAAACTTTTCTTTCTTTTCTTTTTTAAAGTAAACTTGTGTATTGAGTCTTAACTGTATTTCAGTAT
TTTCCAGCCTTATGTGTTACATTATTCCAATGATACCCAACAGTTTATTTTTATTATTTTTTTAAACAAA
ATTTCACAGTTCTGTAATGTAGGCACTTTTATTTTCATTGTGATTTATATATAAGGTAATGTAGGGTTAT
ATTTGGGAGTGACTGCAAGCATTTTTCCATCTGTGTGCAACTAACTGACTCTGTTATTGATCCCTTCTCC
TGCCCTTTCCCAGGTAATTTAAATTGGTCATGGTAGATTTTTTTCATAGATTTGAAAAACTTTTAGGTTG
TTACCAAGTATGAAGTATAAATCTGGGGAAGAGGTTTTATTTACATTTTAGGGTGGGTAAGAAAGCCACC
TTGTTACAAATTTTTTAATTTCCAAAATAATCTATATTAAATGAGGGTTTCTGATCTGTACTTTGTGTTT
AGCTACCTTTTTATATTTAAAAAATTAAAAATGAAAATTACGTTCTTACAAGCTTAAAGCTTGATTTGAT
CTTTGTTTAAATGCCAAAATGTACTTAAATGAGTTACTTAGAATGCCATAAAATTGCAGTTTCATGTATG
TATATAATCATGCTCATGTATATTTAGTTACGTATAATGCTTTCTGAGTGAGTTTTACTCTTAAATCATT
TGGTTAAATCATTTGGCTTGCTGTTTACTCCCTTCTGTAGTTTTTAATTAAAAACTTTAAAGATAAGTCT
ACATTAAACAATGATCACATCTAAAGCTTTATCTTTGTGTAATCTAAGTATATGTGAGAAATCAGAATTG
GCATAATTTGTCTTAGTTGATATTCAAGGCTTTAAAAGTCATTATTCCTGGGCTTGGTAAGTGAATTTAT
GAGATTTACTGCTCTAGAAAGTATAGATGGCGAAAGGACCGTTTTGTATTGCTTCCTGATTACCAGTCTG
ATTATACCATGTGTGCTAATATACTTTTTTTGTTATAGATTGTCTTAATGGTAGGTCAAGTAATAAAAAG
AGATGAAATAATTTAAAAAAAAAAAAAAAA

Human PRKAR1A mRNA sequence - var4 (public gi: 1526989) (SEQ ID NO: 114)
GCTGGGAGCAAAGCGCTGAGGGAGCTCGGTACGCCGCCGCCTCGCACCCGCAGCCTCGCGCCCGCCGCCG
CCCGTCCCCAGAGAACCATGGAGTCTGGCAGTACCGCCGCCAGTGAGGAGGCACGCAGCCTTCGAGAATG

TGAGCTCTACGTCCAGAAGCATAACATTCAAGCGCTGCTCAAAGATTCTATTGTGCAGTTGTGCACTGCT
CGACCTGAGAGACCCATGGCATTCCTCAGGGAATACTTTGAGAGGTTGGAGAAGGAGGAGGCAAAACAGA
TTCAGAATCTGCAGAAAGCAGGCACTCGTACAGACTCAAGGGAGGATGAGATTTCTCCTCCTCCACCCAA
CCCAGTGGTTAAAGGTAGGAGGCGACGAGGTGCTATCAGCGCTGAGGTCTACACGGAGGAAGATGCGGCA
TCCTATGTTAGAAAGGTTATACCAAAAGATTACAAGACAATGGCCGCTTTAGCCAAAGCCATTGAAAAGA
ATGTGCTGTTTTCACATCTTGATGATAATGAGAGAAGTGATATTTTTGATGCCATGTTTTCGGTCTCCTT
TATCGCAGGAGAGACTGTGATTCAGCAAGGTGATGAAGGGGATAACTTCTATGTGATTGATCAAGGAGAG
ACGGATGTCTATGTTAACAATGAATGGGCAACCAGTGTTGGGGAAGGAGGGAGCTTTGGAGAACTTGCTT
TGATTTATGGAACACCGAGAGCAGCCACTGTCAAAGCAAAGACAAATGTGAAATTGTGGGGCATCGACCG
AGACAGCTATAGAAGAATCCTCATGGGAAGCACACTGAGAAAGCGGAAGATGTATGAGGAATTCCTTAGT
AAAGTCTCTATTTTAGAGTCTCTGGACAAGTGGGAACGTCTTACGGTAGCTGATGCATTGGAACCAGTGC
AGTTTGAAGATGGGCAGAAGATTGTGGTGCAGGGAGAACCAGGGGATGAGTTCTTCATTATTTTAGAGGG
GTCAGCTGCTGTGCTACAACGTCGGTCAGAAAATGAAGAGTTTGTTGAAGTGGGAAGATTGGGGCCTTCT
GATTATTTTGGTGAAATTGCACTACTGATGAATCGTCCTCGTGCTGCCACAGTTGTTGCTCGTGGCCCCT
TGAAGTGCGTTAAGCTGGACCGACCTAGATTTGAACGTGTTCTTGGCCCATGCTCAGACATCCTCAAACG
AAACATCCAGCAGTACAACAGTTTTGTGTCACTGTCTGTCTGAAATCTGCCTCCTGTGCCTCCCTTTTCT
CCTCTCCCCAATCCATGCTTCACTCATGCAAACTGCTTTATTTTCCCTACTTGCAGCGCCAAGTGGCCAC
TGGCATCGCAGCTTCCTGTCTGTTTATATATTGAAAGTTGCTTTTATTGCACCATTTTCAATTTGGAGCA
TTAACTAAATGCTCATACACAGTTAAATAAATAGAAAGAGTTCTATGGAGACTTTGCTGTTACTGCTTCT
CTTTGTGCAGTGTTAGTATTCACCCTGGGCAGTGAGTGCCATGCTTTTTGGTGAGGGCAGATCCAGCACC
TATTGAATTACCATAGAGTAATGATGTAACAGTGCAAGATTTTTTTTTTAAGTGACATAATTGTCCAGT
TATAAGCGTATTTAGACTGTGGCCATATATGCTGTATTTCTTTGTAGAATAAATGGTTTCTCATTAAACT
CTAAAGATTAGGGAAATGGATATAGAAAATCTTAGTATAGTAGAAAGACATCTGCCTGTAATTAAACTAG
TTTAAGGGTGGAAAAATGAAAATTTTTGCTAATTATCAATGGGATATGATTGGTTCAGTTTTTTTTTTCC
AGAGTTGTTGTTTGCCAAGCTAATCTGCCTGGTTTATTTATATCTTGTTATTAATGTTTCTTCTCCAATT
CTGAAATACTTTTGAGTATGGCTATCTATACCTGCCTTTTAAGTTTGAAACTAACTCATAGATGCAAATA
TTGGTTAGTATTTAACTACATCTGCCTCGGCTCACAAATTCCGATTAGACCTTTATCCAGCTAGTGCCAA
ATAATTGATCAGATGCTGAATTGAGAATAAGAATTTGAGGTCTACATTCTTGGTTGTTAATTTAGAGCGT
TTGGTTAAAGTATGTCCTTCAGCTGACTCCAGTATAATCTCCTCTGCTCATTAAACTGATTCCAGGAGAT
TGGATTTGCTGTGACTAGATACAGATGGAGCAAATGTCCTAACAGAGAAATAGAGGTGATGCTGCTAAAG
GGAGAAATGCCAGGCGGACAAAGTTCAGTGTCGGGAATTTTCCCCGTGACATTCACTGGGGCATGAGATT
TTGGAAGAAGTTTTTTACTTTGGTTTAGTCTTTTTTTCCTCCTTTTTATTCAGCTAGAATTTCTGGTGGG
TTGATGGTAGGGTATAATGTGTCTGTGTTGCTTCAAATTGGTCTGAAAGGCTATCCTGCTGAAAGTCCTG
CTTTCCTATCTAGCATTTATTCCTCTGGCAAACTTTTCTTTCTTTTCTTTTTTAAAGTAAACTTGTGTAT
TGAGTCTTAACTGTATTTCAGTATTTTCCAGCCTTATGTGTTACATTATTCCAATGATACCCAACAGTTT
ATTTTTATTATTTTTTTAAACAAAATTTCACAGTTCTGTAATGTAGGCACTTTTATTTTCATTGTGATTT
ATATATAAGGTAATGTAGGGTTATATTTGGGAGTGACTGCAAGCATTTTTCCATCTGTGTGCAACTAACT
GACTCTGTTATTGATCCCTTCTCCTGCCCTTTCCCAGGTAATTTAAATTGGTCATGGTAGATTTTTTTCA
TAGATTTGAAAAACTTTTAGGTTGTTACCAAGTATGAAGTATAAATCTGGGGAAGAGGTTTTATTTACAT
TTTAGGGTGGGTAAGAAAGCCACCTTGTTACAAATTTTTTAATTTCCAAAATAATCTATATTAAATGAGG
GTTTCTGATCTGTACTTTGTGTTTAGCTACCTTTTTATATTTAAAAAATTAAAAATGAAAATTATGTTCT
TACAAGCTTAAAGCTTGATTTGATCT

Human PRKAR1A mRNA sequence - var5 (public gi: 1526988) (SEQ ID NO: 115)
GGCAGAGTGGAGCGGGGCTGGAGCAAAGCGCTGAGGGAGCTCGGTACGCCGCCGCCTCGCACCCGCAGC
CTCGCGCCCGCCGCCGCCCGTCCCCAGAGAACCATGGAGTCTGGCAGTACCGCCGCCAGTGAGGAGGCAC
GCAGCCTTCGAGAATGTGAGCTCTACGTCCAGAAGCATAACATTCAAGCGCTGCTCAAAGATTCTATTGT
GCAGTTGTGCACTGCTCGACCTGAGAGACCCATGGCATTCCTCAGGGAATACTTTGAGAGGTTGGAGAAG
GAGGAGGCAAAACAGATTCAGAATCTGCAGAAAGCAGGCACTCGTACAGACTCAAGGGAGGATGAGATTT
CTCCTCCTCCACCCAACCCAGTGGTTAAAGGTAGGAGGCGACGAGGTGCTATCAGCGCTGAGGTCTACAC
GGAGGAAGATGCGGCATCCTATGTTAGAAAGGTTATACCAAAAGATTACAAGACAATGGCCGCTTTAGCC
AAAGCCATTGAAAAGAATGTGCTGTTTTCACATCTTGATGATAATGAGAGAAGTGATATTTTTGATGCCA
TGTTTTCGGTCTCCTTTATCGCAGGAGAGACTGTGATTCAGCAAGGTGATGAAGGGGATAACTTCTATGT
GATTGATCAAGGAGAGACGGATGTCTATGTTAACAATGAATGGGCAACCAGTGTTGGGGAAGGAGGGAGC
TTTGGAGAACTTGCTTTGATTTATGGAACACCGAGAGCAGCCACTGTCAAAGCAAAGACAAATGTGAAAT
TGTGGGGCATCGACCGAGACAGCTATAGAAGAATCCTCATGGGAAGCACACTGAGAAAGCGGAAGATGTA
TGAGGAATTCCTTAGTAAAGTCTCTATTTTAGAGTCTCTGGACAAGTGGGAACGTCTTACGGTAGCTGAT
GCATTGGAACCAGTGCAGTTTGAAGATGGGCAGAAGATTGTGGTGCAGGGAGAACCAGGGGATGAGTTCT
TCATTATTTTAGAGGGGTCAGCTGCTGTGCTACAACGTCGGTCAGAAAATGAAGAGTTTGTTGAAGTGGG
AAGATTGGGGCCTTCTGATTATTTTGGTGAAATTGCACTACTGATGAATCGTCCTCGTGCTGCCACAGTT
GTTGCTCGTGGCCCCTTGAAGTGCGTTAAGCTGGACCGACCTAGATTTGAACGTGTTCTTGGCCCATGCT
CAGACATCCTCAAACGAAACATCCAGCAGTACAACAGTTTTGTGTCACTGTCTGTCTGAAATCTGCCTCC
TGTGCCTCCCTTTTCTCCTCTCCCCAATCCATGCTTCACTCATGCAAACTGCTTTATTTTCCCTACTTGC

AGCGCCAAGTGGCCACTGGCATCGCAGCTTCCTGTCTGTTTATATATTAAAGTTGCTTTTATTGCACCAT
TTTCAATTTGGAGCATTAACTAAATGCTCATACACAGTTAAATAAATAGAAAGAGTTCTATGGAAAAAAA
AAAAAA

Human PRKAR1A mRNA sequence - var6 (public gi: 9956010) (SEQ ID NO: 116)
AACTGACTCTGTTATTGATCCCTTCTCCTGCCCTTTCCCAGGTAATTTAAATTGGTCATGGTAGATTTTT
TTCATAGATTTGAAAAACTTTTAGGTTGTTACCAAGTATGAAGTATAAATCTGGGAAGAGGTTTTATTT
ACATTTTAGGGTGGGTAAGAAAGCCACCTTGTTACAAATTTTTTAATTTCCAAAATAATCTATATTAAAT
GAGGGTTTCTGATCTGTACTTTGTGTTTAGCTACCTTTTTATATTTAAAAAATTAAAAATGAAAATTACG
TTCTTACAAGCTTAAAGCTTGATTTGATCTTTGTTTAAATGCCAAAATGTACTTAAATGAGTTACTTAGA
ATGCCATAAAATTGCAGTTTCATGTATGTATATAATCATGCTCATGTATATTTAGTTACGTATAATGCTT
TCTGAGTGAGTTTTACTCTTAAATCATTTGGTTAAATCATTTGGCTTGCTGTTTACTCCCTTCTGTAGTT
TTTAATTAAAAACTTTAAAGATAAGTCTACATTAAACAATGATCACATCTAAAGCTTTATCTTTGTGTAA
TCTAAGTATATGTGAGAAATCAGAATTGGCATAATTTGTCTTAGTTGATATTCAAGGCTTTAAAAGTCAT
TATTCCTGGGCTTGGTAAGTGAATTTATGAGATTTACTGCTCTAGAAAGTATAGATGGCCAAAGGACCGT
TTTGTATTGCTTCCTGATTACCAGTCTGATTATACCATGTGTGCTAATATACTTTTTTTGTTATAGATTG
TCTTAATGGTAGGTCAAGTAATAAAAAGAGATGAAATAATTTAAATTCTTAAATGAATCAGTTTTTCTTC
CCTTTCTCCTTTCCGTCTTTCCTCTCTCTGTCCTTCCCCGAAAGTCTACTCGGGTGGGCAAAAATGAAAA
GGGGGAAAGTGAATTATGGGATCGGTGTTTTGAAAGAGCAATGTTTATTTTCAGTGCTTTTCAGTTTGTC
AAAGAGTGGATCTCAAAATCTTGCTTAAAGGGTAATTGAGATGTAGCAGATTTATTTACTTAGTCATGGA
AAGAAAAAAAATTCAGTCAAAAGCTAAAGATTTCCTTTTGATTGAAGACAGATTGGTTCTGTGGCCTTGGA
ACTTTCCCAGACTTAATGGGGAAACATCATTTCTAGATTAGCATACTCTTTGGTTTAAATTTAATATATA
CATTTAATGTTACTTAGGGATACTTTTATATTTTGCATATATAAAGCCTCATATATAAAGCCTTATTTCT
GATGCTCTTAGATTTCTGAGGAGTGAGATGATTAAGTTGTATTCATTAGTGTATTGGTATTTCTTCACAT
CCAGTGAAATTGGAGATATGTTGTATGTTAGAAGAGCATTCTTTAAATTGTGTTGCTTTGAACATGTGTA
CCTTTTCTAGATTCAGTAATCCCTTCCCCCCGTCCTCTGGAGTATGAAACCTTTAGAGTCACAATAAAAT
GTAACTAAAGAAAAAAAAAAAAAAAAAAAAAAAAA

Human PRKAR1A mRNA sequence - var7 (public gi: 21757396) (SEQ ID NO: 117)
TAATTTTCTTGTGTGTTTTTAAAAATTTTGATTATGCTAGTAGTTGGCTAATCAGATCCTCACTCCAGTG
GTTTGCTCTGTGACGTTAGGATACTCCCATGGGATAGAAGTTACGTATAGGGAATGTCAGATATTCTTCA
TTGTGCTGACTTGCTTTCGCTTACAGTTGACTTTTGTGCCCTGGTAATTCTGTATCCTGTTTACCGTTTA
CCTACTTCCCACGTCATCATGATTTCTTTTGAGGGAGAACTGAATGAAATTCCCTTAAGGGCCTGACTTC
AGCACCCGTCTCTGCAGAGGTTAGTGGCTCATACTTCCTCCCAGGAGCTGAGGTTATCGACTCTCACTGT
TGCCTACAGAGCACAGATCCTGAACTAAATGAAACATTTACTTGGAATAATGCTAATTCTGTACATATTT
TATTCCCTAGTCCCCACTTCCCTGTTTAAAAACAAAATCTACTTAGAAAAAAATCCCTGTGAATCAGTTG
TCTAATGAATTTAGCAAGTTAAATGCCAGATTGACATTTTGCTTTATAGTTTATACAAGCATGTGTGTGT
TTTTTTCTCGCAGAGAACCATGGAGTCTGGCAGTACCGCCGCCAGTGAGGAGGCACGCAGCCTTCGAGAA
TGTGAGCTCTACGTCCAGAAGCATAACATTCAAGCGCTGCTCAAAGATTCTATTGTGCAGTTGTGCACTG
CTCGACCTGAGAGACCCATGGCATTCCTCAGGGAATACTTTGAGAGGTTGGAGAAGGAGGAGGCAAAACA
GATTCAGAATCTGCAGAAAGCAGGCACTCGTACAGACTCAAGGGAGGATGAGATTTCTCCTCCTCCACCC
AACCCAGTGGTTAAAGGTAGGAGGCGACGAGGTGCTATCAGCGCTGAGGTCTACACGGAGGAAGATGCGG
CATCCTATGTTAGAAAGGTTATACCAAAAGATTACAAGACAATGGCCGCTTTAGCCAAAGCCATTGAAAA
GAATGTGCTGTTTTCACATCTTGATGATAATGAGAGAAGTGATATTTTTGATGCCATGTTTTCGGTCTCC
TTTATCGCAGGAGAGACTGTGATTCAGCAAGGTGATGAAGGGGATAACTTCTATGTGATTGATCAAGGAG
AGACGGATGTCTATGTTAACAATGAATGGGCAACCAGTGTTGGGGAAGGAGGGAGCTTTGGAGAACTTGC
TTTGATTTATGGAACACCGAGAGCAGCCACTGTCAAAGCAAAGACAAATGTGAAATTGTGGGGCATCGAC
CGAGACAGCTATAGAAGAATCCTCATGGGAAGCACACTGAGAAAGCGGAAGATGTATGAGGAATTCCTTA
GTAAAGTCTCTATTTTAGAGTCTCTGGACAAGTGGGAACGTCTTACGGTAGCTGATGCATTGGAACCAGT
GCAGTTTGAAGATGGGCAGAAGATTGTGGTGCAGGGAGAACCAGGGGATGAGTTCTTCATTATTTTAGAG
GGGTCAGCTGCTGTGCTACAACGTCGGTCAGAAAATGAAGAGTTTGTTGAAGTGGGAAGATTGGGGCCTT
CTGATTATTTTGGTGAAATTGCACTACTGATGAATCGTCCTCGTGCTGCCACAGTTGTTGCTCGTGGCCC
CTTGAAGTGCGTTAAGCTGGACCGACCTAGATTTGAACGTGTTCTTGGCCCATGCTCAGACATCCTCAAA
CGAAACATCCAGCAGTACAACAGTTTTGTGTCACTGTCTGTCTGAAATCCGCCTCCTGTGCCTCCCTTTT
CTCCTCTCCCCAATCCATGCTTCACTCATGCAAACTGCTTTATTTTCCCTACTTGCAGCGCCAAGTGGCC
ACTGGCATCGCAGCTTCCTGTCTGTTTATATATTGAAAGTTGCTTTTATTGCACCATTTTCAATTTGGAG
CATTAACTAAATGCTCATACACAGTTAAATAAATAGAAAGAGTTCTATGG

Human PRKAR1A mRNA sequence - var8 (public gi: 1658305) (SEQ ID NO: 118)
AGAGGCGTCAAGGGAGGCCGGAGGGAGAGTGGGGTGGACAGAGGAGCGGAGGGACGAGAGGGAAGCGCAC
GATAGCTGCGCGGAGAGAGAGCGAAGAGCAGGAGGAGGAACAAAGGCGACCCAAGACACCCAGAGAGGGA
CAGAGAACCATGGAGTCTGGCAGTACCGCCGCCAGTGAGGAGGCACGCAGCCTTCGAGAATGTGAGCTCT

ACGTCCAGAAGCATAACATTCAAGCGCTGCTCAAAGATTCTATTGTGCAGTTGTGCACTGCTCGACCTGA
GAGACCCATGGCATTCCTCAGGGAATACTTTGAGAGGTTGGAGAAGGAGGAGGCAAAACAGATTCAGAAT
CTGCAGAAAGCAGGCACTCGTACAGACTCAAGGGAGGATGAGATTTCTCCTCCTCCACCCAA

Human PRKAR1A protein sequence - var1 (public gi: 4506063) (SEQ ID NO: 264)
MESGSTAASEEARSLRECELYVQKHNIQALLKDSIVQLCTARPERPMAFLREYFERLEKEEAKQIQNLQK
AGTRTDSREDEISPPPPNPVVKGRRRRGAISAEVYTEEDAASYVRKVIPKDYKTMAALAKAIEKNVLFSH
LDDNERSDIFDAMFSVSFIAGETVIQQGDEGDNFYVIDQGETDVYVNNEWATSVGEGGSFGELALIYGTP
RAATVKAKTNVKLWGIDRDSYRRILMGSTLRKRKMYEEFLSKVSILESLDKWERLTVADALEPVQFEDGQ
KIVVQGEPGDEFFIILEGSAAVLQRRSENEEFVEVGRLGPSDYFGEIALLMNRPRAATVVARGPLKCVKL
DRPRFERVLGPCSDILKRNIQQYNSFVSLSV

Human PRKAR1A protein sequence - var2 (public gi: 1658306) (SEQ ID NO: 265)
MESGSTAASEEARSLRECELYVQKHNIQALLKDSIVQLCTARPERPMAFLREYFERLEKEEAKQIQNLQK
AGTRTDSREDEISPPPP

Human PRKAR1A pray sequence - var1 (SEQ ID NO: 119)
GCCGCCTGGTNTACCCATACGACGTACCAGTATTACGCTCATATGGCCATGGCAGGCCAGTGCAATTCCA
CCCAAGCAGTGGCTATCAACGCAGAGTGGTAGCGGGGCATGGGAGCAAAGCAGCATGAGGGAGCTCGGTA
CNCCGCCGCCTCNCACCCGCAGCCTCGCGCCCGCCGCCGCCCGTCCCCAGNGAACCATGGAGTCTGGCAG
TACCGTTTCCAGTGAGGAGGCACNCAGCCTTCGAGAATGTGAGCTCTNNGTCCAGAAGCATNACATTCAN
TGCGCTNCTCAAAGATTCTNTTGTGCANTTGTGCNCTGCTCGACCTNAGAGACCGGGTGGCATTCCTCAN
GGAATACTTGGCGNACGNNGNNTAATGANGAGGCCCNNTNTNTTNCAAANTCTNCANNTNTTTNNNTCTT
TNACAAACTTTTTGGACNATNANNNANCCCNTNNNANANAAANAANAATNNCTTCCCCGGGGGNATTCCT
NCCC

Human PRKAR1A pray sequence - var2 (SEQ ID NO: 120)
GAGCGCCGCCATGGNANTACCCATACGACGTACCAGNATTACGCTCATATGGCCATGGAGGCCAGTGAAT
TCCACCCAAGCAGTGGTATCAACGCAGAGTGGTAGCGGGGCTGGGAGCAAAGCGCTGAGGGAGCTCGGTA
CGCCGCCGCCTCGCACCCGCAGCCTCGCGCCCGCCGCCGCCCGTCCCCAGAGAACCATGGAGTCTGGCAG
TACCGCCGCCAGTGAGGAGGCACGCAGCCTTCGAGAATGTGAGCTCTACGTCCAGAAGCATAACATTCAA
GCGCTGCTCAAAGATTCTATTGTGCAGTTGTGCACTGCTCGACCTGAGAGACCCATGGCATTCCTCAGGG
AATACTTTGAGAGGTTGGAGAAGGAGGAGGCAAAACAGATTCAGAATCTGCAGAAAGCAGGCACTCGTAC
AGACTCAAGGGAGGATGAGATTTCTCCTCCTCCACCCAACCCAGTGGTTAAAGGTAGGAGGCGACGAGGT
GCTATCAGCGCTGAGGTCTACACGGAGGAAGATGCGGCATCCTATGTTAGAAAGGTTATACCAAAAGATT
ACAAGACGATGGCCGCTTTAGCCAAAGCCATTGAAAAGAATGTGCTGTTTTCACATCTTGATGATAATGA
GAGAAGTGATATTTTTGATGCCATGTTTTCGGTCTCCTTTATCGCAGGAGAGACTGTGATTCANCAAGGT
GATGAAGGGGATAACTTCTATGTGATTGATCAAGGANAGACNGATGTCTATGTTAACAATGAATGGGCNA
CCANTGTTGGGGAAGGAGGGAGCTTTGGAAAACTTGCTTTGATTNANGGAANCCNNNNGCNNCCNTNGTC
AAACCAAAACAAA

Human PRKAR1A pray sequence - var3 (SEQ ID NO: 121)
CGACGCCGCCTGGTATACCCATACGACGTACCAGTATTACGCTCATATGGCCATGGCAGGCCAGTGAATT
CCACCCAAGCAGGTGCGATATGCATACGCGAGNAGTGAGTAACGGCGGCTGGGTAGCGAAGTCGCTGAGG
GAGCTCGGTACNCCGCCAGCGCTCGCACCCGCANCCTCGCGCCCGCCGCCGCCCGTCCCCAGAGAACCAT
GGAGTCTGGCAGTACCGCCGCCAGTGAGGAGGCACGCAGCCTTCGAGAATGTGAGCTCTACGTCCAGAAG
CATAACATTCAAGCGCTGCTCAAAGATTCTATTGTGCAGTTGTGCACTGCTCGACCTGAGAGACCCATGG
CATTCCTCAGGGAATACTTTGAGAGGTTGGAGAAGGAGGAGGCAAAACAGATTCAGAATCTGCAGAAAGC
AGGCACTCGTACAGACTCAAGGGAGGATGAGATTTCTCCTCCTCCACCCAACCCAGTGGTTAAAGGTAGG
AGGCGACGAGGTGCTATCAGCGCTGAGGTCTACACGGAGGAAGATGCGGCATCCTATGTTAGAAAGGTAG
TTTTTGATATTTGAATATCGGGGGGGATGCTTTNGGGACCCACTTGGTGGTCATCTANTCTCCTTGGATG
ANTGATTCTTAAATCCAAAACNGGGNGGAACTTCATCNNNCTTNTANANTNNTGGGNNCTGGAAAAANGG
TTTTTNTAATACCNNCTTNNCAANGAAANANCNNTTNGNGTTTTNAANNNNGGAAAANTGGCTTTNGGGG
GTTNNNNTTTCCNTCNNNNTNTTTNNNNAAAAAGGGNGGGGGCGGTTNG

Human PRKAR1A pray sequence - var4 (SEQ ID NO: 122)
CGTANCNNCGCGNGACTCGGTGACTGANGCCATGATCGCACATTACACACTATNTACCGTCTGACATCAT
GGNTCAGTGTGCAGGGCCATGTTGANNTCTCCNCNCCATANATACAAGGNCTCAAGNNGNACANAACAAT
AGAGANATATTTTTANTACTNACTCACTATAGGGCGAGCGCCGCCATGGAGTACCCATACGACGTNCCAG
ATTACGCTCATATGGCCATGGAGGCCAGTGAATTCCACCCAAGCAGTGGTATCAACGCAGAGTGGAGCGG
GGCTGGGAGCAAAGCGCTGAGGGAGCTCGGTACGCCGCCGCCTCGCACCCGCAGCCTCGCGCCCGCCGCC

GCCCGTCCCCAGAGAACCATGGAGTCTGGCNGTACCGCCNNTANTGNGGAGGCACGCAGCCTTNNAGAAT
GTGAGCTCTACGTCCAGAAGCATAACATNNGNGCGCTGCTCAAAGATTCTATTGTGCAGTTGTGCACTGC
TCGACCTGAGAGACCCATGGCATTCCTCAGGGAATTACTTTGAGAGGTTGGANNAGGAGGAGGCNAACCA
NATTCANAATCTGCNGAAAGCANNANTCNTACAGACTCAGGGGNGGNNANATTTTTTATTCTTCCCCCCA
NCCNANTGGTTAAGGGTNGGAGGCNACAAGGNCTNTTNNCCCCTGAAGGNNTNCCCGGNGGAAGATNCGG
ATTCCTATGTTAAAANGGGTNTTTCCNNTANNNATTNCNANNAANANGGCCCCTTTTNNCCCAAANCCCT
TCNAAAAAAANGNGCNNTTTCCNANTNTNNGNGAANTTNNAAAAAGNGGNTTTTTTTTAAANCCCNTTTT
TNNCGTTNTCTTTTTCNGGNGGAAACNTTNATTAANNCCG

Unigene Name: PRKARIA Unigene ID: Hs.183037 Clone ID: 3GD_188

Human PRKARIA mRNA sequence - var1 (public gi: 23273779) (SEQ ID NO: 396)
GGTGGAGCTGTCGCCTAGCCGCTATCGCAGAGTGGAGCGGGGCTGGGAGCAAAGCGCTGAGGGAGCTCGG
TACGCCGCCGCCTCGCACCCGCAGCCTCGCGCCCGCCGCCGCCCGTCCCCAGAGAACCATGGAGTCTGGC
AGTACCGCCGCCAGTGAGGAGGCACGCAGCCTTCGAGAATGTGAGCTCTACGTCCAGAAGCATAACATTC
AAGCGCTGCTCAAAGATTCTATTGTGCAGTTGTGCACTGCTCGACCTGAGAGACCCATGGCATTCCTCAG
GGAATACTTTGAGAGGTTGGAGAAGGAGGAGGCAAAACAGATTCAGAATCTGCAGAAAGCAGGCACTCGT
ACAGACTCAAGGGAGGATGAGATTTCTCCTCCTCCACCCAACCCAGTGGTTAAAGGTAGGAGGCGACGAG
GTGCTATCAGCGCTGAGGTCTACACGGAGGAAGATGCGGCATCCTATGTTAGAAAGGTTATACCAAAAGA
TTACAAGACAATGGCCGCTTTAGCCAAAGCCATTGAAAAGAATGTGCTGTTTTCACATCTTGATGATAAT
GAGAGAAGTGATATTTTTGATGCCATGTTTTCGGTCTCCTTTATCGCAGGAGAGACTGTGATTCAGCAAG
GTGATGAAGGGGATAACTTCTATGTGATTGATCAAGGAGAGACGGATGTCTATGTTAACAATGAATGGGC
AACCAGTGTTGGGGAAGGAGGGAGCTTTGGAGAACTTGCTTTGATTTATGGAACACCGAGAGCAGCCACT
GTCAAAGCAAAGACAAATGTGAAATTGTGGGCATCGACCGAGACAGCTATAGAAGAATCCTCATGGGAA
GCACACTGAGAAAGCGGAAGATGTATGAGGAATTCCTTAGTAAAGTCTCTATTTTAGAGTCTCTGGACAA
GTGGGAACGTCTTACGGTAGCTGATGCATTGGAACCAGTGCAGTTTGAAGATGGGCAGAAGATTGTGGTG
CAGGGAGAACCAGGGGATGAGTTCTTCATTATTTTAGAGGGGTCAGCTGCTGTGCTACAACGTCGGTCAG
AAAATGAAGAGTTTGTTGAAGTGGGAAGATTGGGGCCTTCTGATTATTTTGGTGAAATTGCACTACTGAT
GAATCGTCCTCGTGCTGCCACAGTTGTTGCTCGTGGCCCCTTGAAGTGCGTTAAGCTGGACCGACCTAGA
TTTGAACGTGTTCTTGGCCCATGCTCAGACATCCTCAAACGAAACATCCAGCAGTACAACAGTTTTGTGT
CACTGTCTGTCTGAAATCTGCCTCCTGTGCCTCCCTTTTCTCCTCTCCCCAATCCATGCTTCACTCATGC
AAACTGCTTTATTTTCCCTACTTGCAGCGCCAAGTGGCCACTGGCATCGCAGCTTCCTGTCTGTTTATAT
ATTGAAAGTTGCTTTTATTGCACCATTTTCAATTTGGAGCATTAACTAAATGCTCATACACAGTTAAATA
AATAGAAAGAGTTCTATGGAGACTTTGCTGTTACTGCTTCTCTTTGTGCAGTGTTAGTATTCACCCTGGG
CAGTGAGTGCCATGCTTTTTGGTGAGGGCAGATCCCAGCACCTATTGAATTACCATAGAGTAATGATGTA
ACAGTGCAAGATTTTTTTTTAAGTGACATAATTGTCCAGTTATAAGCGTATTTAGACTGTGGCCATATA
TGCTGTATTTCTTTGTAGAATAAATGGTTTCTCATTAAACTCTAAAGATTAGGGAAAATGGATATAGAAA
ATCTTAGTATAGTAGAAAGACATCTGCCTGTAATTAAACTAGTTTAAGGGTGGAAAAATGCCCATTTTTG
CTAATTATCAATGGGATATGATTGGTTCAGTTTTTTTTTTTCCAGAGTTGTTGTTTGCCAAGCTAATCTG
CCTGGTTTTATTTATATCTTGTTATTAATGTTTCTTCTCCAATTCTGAAATACTTTTGAGTATGGCTATC
TATACCTGCCTTTTAAGTTTGAAACTAACTCATAGATTGCAAATATTGGTTAGTATTTAACTACATCTGC
CTCGGCTCACAAATTCCGATTAGACCTTTATCCAGCTAGTGCCAAATAATTGATCAGATGCTGAATTGAG
AATAAGAATTTGAGGTCTACATTCTTGGTTGTTAATTTAGAGCGTTTGGTTAAAGTATGTCCTTCAGCTG
ACTCCAGTATAATCTCCTCTGCTCATTAAACTGATTCCAGGAGATTGGATTTGCTGTGACTAGATACAGA
TGGAGCAAATGTCCTAACAGAGAAATAGAGGTGATGCTGCTAAAGGGAGAAATGCCAGGCGGACAAAGTT
CAGTGTCGGGAATTTTCCCCGTGACATTCACTGGGGCATGAGATTTTGGAAGAAGTTTTTTACTTTGGTT
TAGTCTTTTTTTCCTTCCTTTTTATTCAGCTAGAATTTCTGGTGGGTTGATGGTAGGGTATAATGTGTCT
GTGTTGCTTCAAATTGGTCTGAAAGGCTATCCTGCGGAAAGTCCTGCTTTCCTATCTAGCATTTATTTCT
CTGGCAAACTTTTCTTTCTTTTCTTTTTTAAAGTAAACTTGTGTATTGAGTCTTAACTGTATTTCAGTAT
TTTCCAGCCTTATGTGTTACATTATTCCAATGATACCCAACAGTTTATTTTTATTATTTTTTTAAACAAA
ATTTCACAGTTCTGTAATGTAGGCACTTTTATTTTCATTGTGATTTATATATAAGGTAATGTAGGGTTAT
ATTTGGAGTGACTGCAAGCATTTTTCCATCTGTGTGCAACTAACTGACTCTGTTATTGATCCCTTCTCC
TGCCCTTTCCCAGGTAATTTAAATTGGTCATGGTAGATTTTTTTCATAGATTTGAAAAACTTTTAGGTTG
TTACCAAGTATGAAGTATAAATCTGGGGAAGAGGTTTTATTTACATTTTAGGGTGGGTAAGAAAGCCACC
TTGTTACAAATTTTTTAATTTCCAAAATAATCTATATTAAATGAGGGTTTCTGATCTGTACTTTGTGTTT
AGCTACCTTTTTATATTTAAAAAATTAAAAATGAAAATTACGTTCTTACAAGCTTAAAGCTTGATTTGAT
CTTTGTTTAAATGCCAAAATGTACTTAAATGAGTTACTTAGAATGCCATAAAATTGCAGTTTCATGTATG

TATATAATCATGCTCATGTATATTTAGTTACGTATAATGCTTTCTGAGTGAGTTTTACTCTTAAATCATT
TGGTTAAATCATTTGGCTTGCTGTTTACTCCCTTCTGTAGTTTTTAATTAAAAACTTTAAAGATAAGTCT
ACATTAAACAATGATCACATCTAAAGCTTTATCTTTGTGTAATCTAAGTATATGTGAGAAATCAGAATTG
GCATAATTTGTCTTAGTTGATATTCAAGGCTTTAAAAGTCATTATTCCTGGGCTTGGTAAGTGAATTTAT
GAGATTTACTGCTCTAGAAAGTATAGATGGCGAAAGGACCGTTTTGTATTGCTTCCTGATTACCAGTCTG
ATTATACCATGTGTGCTAATATACTTTTTTTGTTATAGATTGTCTTAATGGTAGGTCAAGTAATAAAAAG
AGATGAAATAATTTAAAAAAAAAAAAAAAA

Human PRKARIA mRNA sequence - (public gi: 4506062) (SEQ ID NO: 397)
GCTGGGAGCAAAGCGCTGAGGGAGCTCGGTACGCCGCCGCCTCGCACCCGCAGCCTCGCGCCCGCCGCCG
CCCGTCCCCAGAGAACCATGGAGTCTGGCAGTACCGCCGCCAGTGAGGAGGCACGCAGCCTTCGAGAATG
TGAGCTCTACGTCCAGAAGCATAACATTCAAGCGCTGCTCAAAGATTCTATTGTGCAGTTGTGCACTGCT
CGACCTGAGAGACCCATGGCATTCCTCAGGGAATACTTTGAGAGGTTGGAGAAGGAGGAGGCAAAACAGA
TTCAGAATCTGCAGAAAGCAGGCACTCGTACAGACTCAAGGGAGGATGAGATTTCTCCTCCTCCACCCAA
CCCAGTGGTTAAAGGTAGGAGGCGACGAGGTGCTATCAGCGCTGAGGTCTACACGGAGGAAGATGCGGCA
TCCTATGTTAGAAAGGTTATACCAAAAGATTACAAGACAATGGCCGCTTTAGCCAAAGCCATTGAAAAGA
ATGTGCTGTTTTCACATCTTGATGATAATGAGAGAAGTGATATTTTTGATGCCATGTTTTCGGTCTCCTT
TATCGCAGGAGAGACTGTGATTCAGCAAGGTGATGAAGGGGATAACTTCTATGTGATTGATCAAGGAGAG
ACGGATGTCTATGTTAACAATGAATGGGCAACCAGTGTTGGGGAAGGAGGGAGCTTTGGAGAACTTGCTT
TGATTTATGGAACACCGAGAGCAGCCACTGTCAAAGCAAAGACAAATGTGAAATTGTGGGGCATCGACCG
AGACAGCTATAGAAGAATCCTCATGGGAAGCACACTGAGAAAGCGGAAGATGTATGAGGAATTCCTTAGT
AAAGTCTCTATTTTAGAGTCTCTGGACAAGTGGGAACGTCTTACGGTAGCTGATGCATTGGAACCAGTGC
AGTTTGAAGATGGGCAGAAGATTGTGGTGCAGGGAGAACCAGGGGATGAGTTCTTCATTATTTTAGAGGG
GTCAGCTGCTGTGCTACAACGTCGGTCAGAAAATGAAGAGTTTGTTGAAGTGGGAAGATTGGGGCCTTCT
GATTATTTTGGTGAAATTGCACTACTGATGAATCGTCCTCGTGCTGCCACAGTTGTTGCTCGTGGCCCCT
TGAAGTGCGTTAAGCTGGACCGACCTAGATTTGAACGTGTTCTTGGCCCATGCTCAGACATCCTCAAACG
AAACATCCAGCAGTACAACAGTTTTGTGTCACTGTCTGTCTGAAATCTGCCTCCTGTGCCTCCCTTTTCT
CCTCTCCCCAATCCATGCTTCACTCATGCAAACTGCTTTATTTTCCCTACTTGCAGCGCCAAGTGGCCAC
TGGCATCGCAGCTTCCTGTCTGTTTATATATTGAAAGTTGCTTTTATTGCACCATTTTCAATTTGGAGCA
TTAACTAAATGCTCATACACAGTTAAATAAATAGAAAGAGTTCTATGGAGACTTTGCTGTTACTGCTTCT
CTTTGTGCAGTGTTAGTATTCACCCTGGGCAGTGAGTGCCATGCTTTTTGGTGAGGGCAGATCCAGCACC
TATTGAATTACCATAGAGTAATGATGTAACAGTGCAAGATTTTTTTTTTTAAGTGACATAATTGTCCAGT
TATAAGCGTATTTAGACTGTGGCCATATATGCTGTATTTCTTTGTAGAATAAATGGTTTCTCATTAAACT
CTAAAGATTAGGGAAATGGATATAGAAAATCTTAGTATAGTAGAAAGACATCTGCCTGTAATTAAACTAG
TTTAAGGGTGGAAAAATGAAAATTTTTGCTAATTATCAATGGGATATGATTGGTTCAGTTTTTTTTTTCC
AGAGTTGTTGTTTGCCAAGCTAATCTGCCTGGTTTATTTATATCTTGTTATTAATGTTTCTTCTCCAATT
CTGAAATACTTTTGAGTATGGCTATCTATACCTGCCTTTTAAGTTTGAAACTAACTCATAGATGCAAATA
TTGGTTAGTATTTAACTACATCTGCCTCGGCTCACAAATTCCGATTAGACCTTTATCCAGCTAGTGCCAA
ATAATTGATCAGATGCTGAATTGAGAATAAGAATTTGAGGTCTACATTCTTGGTTGTTAATTTAGAGCGT
TTGGTTAAAGTATGTCCTTCAGCTGACTCCAGTATAATCTCCTCTGCTCATTAAACTGATTCCAGGAGAT
TGGATTTGCTGTGACTAGATACAGATGGAGCAAATGTCCTAACAGAGAAATAGAGGTGATGCTGCTAAAG
GGAGAAATGCCAGGCGGACAAAGTTCAGTGTCGGGAATTTTCCCCGTGACATTCACTGGGGCATGAGATT
TTGGAAGAAGTTTTTTACTTTGGTTTAGTCTTTTTTTCCTCCTTTTTATTCAGCTAGAATTTCTGGTGGG
TTGATGGTAGGGTATAATGTGTCTGTGTTGCTTCAAATTGGTCTGAAAGGCTATCCTGCTGAAAGTCCTG
CTTTCCTATCTAGCATTTATTCCTCTGGCAAACTTTTCTTTCTTTTCTTTTTTAAAGTAAACTTGTGTAT
TGAGTCTTAACTGTATTTCAGTATTTTCCAGCCTTATGTGTTACATTATTCCAATGATACCCAACAGTTT
ATTTTTATTATTTTTTTAAACAAAATTTCACAGTTCTGTAATGTAGGCACTTTTATTTTCATTGTGATTT
ATATATAAGGTAATGTAGGGTTATATTTGGGAGTGACTGCAAGCATTTTTCCATCTGTGTGCAACTAACT
GACTCTGTTATTGATCCCTTCTCCTGCCCTTTCCCAGGTAATTTAAATTGGTCATGGTAGATTTTTTTCA
TAGATTTGAAAAACTTTTAGGTTGTTACCAAGTATGAAGTATAAATCTGGGGAAGAGGTTTTATTTACAT
TTTAGGGTGGGTAAGAAAGCCACCTTGTTACAAATTTTTTAATTTCCAAAATAATCTATATTAAATGAGG
GTTTCTGATCTGTACTTTGTGTTTAGCTACCTTTTTATATTTAAAAAATTAAAAATGAAAATTATGTTCT
TACAAGCTTAAAGCTTGATTTGATCT

<u>Unigene Name:</u> PTPN12 <u>Unigene ID:</u> Hs.62

Human PTPN12 mRNA sequence - var1 (public gi: 292408) (SEQ ID NO: 123)
AGCGACCGCAGCCGGGGGACGCGGGAGGATGGAGCAAGTGGAGATCCTGAGGAAATTCATCCAGAGGGT
CCAGGCCATGAAGAGTCCTGACCACAATGGGGAGGACAACTTCGCCCGGGACTTCATGCGGTTAAGAAGA
TTGTCTACCAAATATAGAACAGAAAAGATATATCCCACAGCCACTGGAGAAAAAGAAGAAAATGTTAAAA
AGAACAGATACAAGGACATACTGCCATTTGATCACAGCCGAGTTAAATTGACATTAAAGACTCCTTCACA
AGATTCAGACTATATCAATGCAAATTTTATAAAGGGCGTCTATGGCCAAAAGCATATGTAGCAACTCAA
GGACCTTTAGCAAATACAGTAATAGATTTTTGGAGGATGATATGGAGTATAATGTTGTGATCATTGTAA
TGGCCTGCCGAGAATTTGAGATGGGAAGGAAAAAATGTGAGCGCTATTGGCCTTTGTATGGAGAAGACCC
CATAACGTTTGCACCATTTAAAATTTCTTGTGAGGATGAACAAGCAAGAACAGACTACTTCATCAGGACA
CTCTTACTTGAATTTCAAAATGAATCTCGTAGGCTGTATCAGTTTCATTATGTGAACTGGCCAGACCATG
ATGTTCCTTCATCATTTGATTCTATTCTGGACATGATAAGCTTAATGAGGAAATATCAAGAACATGAAGA
TGTTCCTATTTGTATTCATTGCAGTGCAGGCTGTGGAAGAACAGGTGCCATTTGTGCCATAGATTATACG
TGGAATTTACTAAAAGCTGGGAAAATACCAGAGGAATTTAATGTATTTAATTTAATACAAGAAATGAGAA
CACAAAGGCATTCTGCAGTACAAACAAAGGAGCAATATGAACTTGTTCATAGAGCTATTGCCCAACTGTT
TGAAAAACAGCTACAACTATATGAAATTCATGGAGCTCAGAAAATTGCTGATGGAGTGAATGAAATTAAC
ACTGAAAACATGATCAGCTCCATAGAGCCTGAAAAACAAGATTCTCCTCCTCCAAAACCACCAAGGACCC
GCAGTTGCCTTGTTGAAGGGGATGCTAAAGAAGAAATACTGCAGCCACCGGAACCTCATCCAGTGCCACC
CATCTTGACACCTTCTCCCCCTTCAGCTTTTCCAACAGTCACTACTGTGTGGCAGGACAATGATAGATAC
CATCCAAAGCCAGTGTTGCATATGGTTTCATCAGAACAACATTCAGCAGACCTCAACAGAAACTATAGTA
AATCAACAGAACTTCCAGGGAAAAATGAATCAACAATTGAACAGATAGATAAAAAATTGGAACGAAATTT
AAGTTTTGAGATTAAGAAGGTCCCTCTCCAAGAGGGACCAAAAAGTTTTGATGGAACACACTTTTGAAT
AGGGGACATGCAATTAAAATTAAATCTGCTTCACCTTGTATAGCTGATAAAATCTCTAAGCCACAGGAAT
TAAGTTCAGATCTAAATGTCGGTGATACTTCCCAGAATTCTTGTGTGGACTGCAGTGTAACACAATCAAA
CAAAGTTTCAGTTACTCCACCAGAAGAATCCCAGAATTCAGACACACCTCCAAGGCCAGACCGCTTGCCT
CTTGATGAGAAAGGACATGTAACGTGGTCATTTCATGGACCTGAAAATGCCATACCCATACCTGATTTAT
CTGAAGGCAATTCCTCAGATATCAACTATCAAACTAGGAAAACTGTGAGTTTAACACCAAGTCCTACAAC
ACAAGTTGAAACACCTGATCTTGTGGATCATGATAACACTTCACCACTCTTCAGAACACCCCTCAGTTTT
ACTAATCCACTTCACTCTGATGACTCAGACTCAGATGAAAGAAACTCTGATGGTGCTGTGACCCAGAATA
AAACTAATATTTCAACAGCAAGTGCCACAGTTTCTGCTGCCACTAGTACTGAAAGCATTTCTACTAGGAA
AGTATTGCCAATGTCCATTGCTAGACATAATATAGCAGGAACAACACATTCAGGTGCTGAAAAAGATGTT
GATGTTAGTGAAGATTCACCTCCTCCCCTACCTGAAAGAACTCCTGAATCGTTTGTGTTAGCAAGTGAAC
ATAATACACCTGTAAGATCGGAATGGAGTGAACTTCAAAGTCAGGAACGATCTGAACAAAAAAAGTCTGA
AGGCTTGATAACCTCTGAAAATGAGAAATGTGATCATCCAGCGGGAGGTATTCACTATGAAATGTGCATA
GAATGTCCACCTACTTTCAGTGACAAGAGAGAACAAATATCAGAAAATCCAACAGAAGCCACAGATATTG
GTTTTGGTAATCGATGTGGAAAACCCAAAGGACCAAGAGATCCACCTTCAGAATGGACATGATTCAGGGA
GCTAGAAGACACTTTAAGTTATACTGGAAAATTCAGGTGCCACTGAAAGCCAGATTTATAGTATTCCATC
TTTAATATGTGGGACTAACAGCAGTGTAGATTGTTACCTTAATATTTTTTGCTGGGACCATCTACCTGCC
TTATACTACACTTAGGAAAAAGTATTACATATGGTTTATTTTGAAACTTCAAGTATTATTGCCTTAATGT
CTCTTAACCCTGTTACACGCTGCTTGTAGACATGTTAATATAGTAATACCTTTATGATATATTGAGTTTA
AGGACTACTCTTTTTCTGTTTTATCATGTATGCATTATTTTGTATATGTACAGGGCAAGTAGGTATATAA
TTTGATAAAGTTGCAATTGAAATATTATTAACAGAAGATGTAAGAAATTTCTGCATGGTCTAAATCTTTG
TGTACTTTATTTGTAAATTATTTGCCCTGGAGTTTTAGAAAATAGTTTCTGAATTTTAAACTTGCTGGAT
TCATGCAGCCAGCTTTGCAGGTTATCAGAGATCAAAGATTGTAATAATAATTTTGTAAATTGTAAGCAAA
AAGTTATTTTTATATTATATACAGTCTAATTGTTCATCCTAATTGTTCCTGTTTTCATCTAGTCAGAGAT
TCAGTAAGTGCCTTGGAACAATATTGAATTCTCTTAGCTTGTGTGTGTTTCTTTAATATTTGAACTCAAG
TGGGATTAGAAGACTATCAAAATACATGTATGTTTCAGATATTTGACCTGTCATTAAAAAAAACAAACAG
TTTTACAGTG

Human PTPN12 mRNA sequence - var2 (public gi: 29476876) (SEQ ID NO: 124)
GGGGAGAGGCGGCTGCGGCTGCGGCTGCGGCTGCTGGCGGGGGTGGGGGGAGGAGGAACCGGGAAGGG
GGGGCAGGGCGAGCGGAGAGCTAGCTGTGTTCCTGAGGCGCAGCCGCCGCCTAGGGCGGTGGGGAGGAGG
AGGGAGCCGCGGGGCTTGGCGGGGTCGGAGGGAGGGACGTGCTGGGGGAACGAGCTGGGGAAGACGGAG
CGGGCTCTGTGCCGGGCGGGCGGGCGGCGGGGGGGCCAGCGACCGCAGCCGGGGGACGCGGGAGGATGG
AGCAAGTGGAGATCCTGAGGAAATTCATCCAGAGGGTCCAGGCCATGAAGAGTCCTGACCACAATGGGGA
GGACAACTTCGCCCGGGACTTCATGCGGTTAAGAAGATTGTCTACCAAATATAGAACAGAAAAGATATAT
CCCACAGCCACTGGAGAAAAAGAAGAAAATGTTAAAAAGAACAGATACAAGGACATACTGCCATTTGATC
ACAGCCGAGTTAAATTGACATTAAAGACTCCTTCACAAGATTCAGACTATATCAATGCAAATTTTATAAA
GGGCGTCTATGGCCAAAAGCATATGTAGCAACTCAAGGACCTTTAGCAAATACAGTAATAGATTTTTGG
AGGATGATATGGAGTATAATGTTGTGATCATTGTAATGGCCTGCCGAGAATTTGAGATGGGAAGGAAAA
AATGTGAGCGCTATTGGCCTTTGTATGGAGAAGACCCCATAACGTTTGCACCATTTAAAATTTCTTGTGA
GGATGAACAAGCAAGAACAGACTACTTCATCAGGACACTCTTACTTGAATTTCAAAATGAATCTCGTAGG
CTGTATCAGTTTCATTATGTGAACTGGCCAGACCATGATGTTCCTTCATCATTTGATTCTATTCTGGACA

TGATAAGCTTAATGAGGAAATATCAAGAACATGAAGATGTTCCTATTTGTATTCATTGCAGTGCAGGCTG
TGGAAGAACAGGTGCCATTTGTGCCATAGATTATACGTGGAATTTACTAAAAGCTGGGAAAATACCAGAG
GAATTTAATGTATTTAATTTAATACAAGAAATGAGAACACAAAGGCATTCTGCAGTACAAACAAAGGAGC
AATATGAACTTGTTCATAGAGCTATTGCCCAACTGTTTGAAAAACAGCTACAACTATATGAAATTCATGG
AGCTCAGAAAATTGCTGATGGAGTGAATGAAATTAACACTGAAAACATGATCAGCTCCATAGAGCCTGAA
AAACAAGATTCTCCTCCTCCAAAACCACCAAGGACCCGCAGTTGCCTTGTTGAAGGGGATGCTAAAGAAG
AAATACTGCAGCCACCGGAACCTCATCCAGTGCCACCCATCTTGACACCTTCTCCCCCTTCAGCTTTTCC
AACAGTCACTACTGTGTGGCAGGACAATGATAGATACCATCCAAAGCCAGTGTTGCATATGGTTTCATCA
GAACAACATTCAGCAGACCTCAACAGAAACTATAGTAAATCAACAGAACTTCCAGGGAAAAATGAATCAA
CAATTGAACAGATAGATAAAAAATTGGAACGAAATTTAAGTTTTGAGATTAAGAAGGTCCCTCTCCAAGA
GGGACCAAAAAGTTTTGATGGGAACACACTTTTGAATAGGGGACATGCAATTAAAATTAAATCTGCTTCA
CCTTGTATAGCTGATAAAATCTCTAAGCCACAGGAATTAAGTTCAGATCTAAATGTCGGTGATACTTCCC
AGAATTCTTGTGTGGACTGCAGTGTAACACAATCAAACAAAGTTTCAGTTACTCCACCAGAAGAATCCCA
GAATTCAGACACACCTCCAAGGCCAGACCGCTTGCCTCTTGATGAGAAAGGACATGTAACGTGGTCATTT
CATGGACCTGAAAATGCCATACCCATACCTGATTTATCTGAAGGCAATTCCTCAGATATCAACTATCAAA
CTAGGAAAACTGTGAGTTTAACACCAAGTCCTACAACACAAGTTGAAACACCTGATCTTGTGGATCATGA
TAACACTTCACCACTCTTCAGAACACCCCTCAGTTTTACTAATCCACTTCACTCTGATGACTCAGACTCA
GATGAAAGAAACTCTGATGGTGCTGTGACCCAGAATAAAACTAATATTTCAACAGCAAGTGCCACAGTTT
CTGCTGCCACTAGTACTGAAAGCATTTCTACTAGGAAAGTATTGCCAATGTCCATTGCTAGACATAATAT
AGCAGGAACAACACATTCAGGTGCTGAAAAAGATGTTGATGTTAGTGAAGATTCACCTCCTCCCCTACCT
GAAAGAACTCCTGAATCGTTTGTGTTAGCAAGTGAACATAATACACCTGTAAGATCGGAATGGAGTGAAC
TTCAAAGTCAGGAACGATCTGAACAAAAAAAGTCTGAAGGCTTGATAACCTCTGAAAATGAGAAATGTGA
TCATCCAGCGGGAGGTATTCACTATGAAATGTGCATAGAATGTCCACCTACTTTCAGTGACAAGAGAGAA
CAAATATCAGAAAATCCAACAGAAGCCACAGATATTGGTTTTGGTAATCGATGTGGAAAACCCAAAGGAC
CAAGAGATCCACCTTCAGAATGGACATGATTCAGGGAGCTAGAAGACACTTTAAGTTATACTGGAAAATT
CAGGTGCCACTGAAAGCCAGATTTATAGTATTCCATCTTTAATATGTGGGACTAACAGCAGTGTAGATTG
TTACCTTAATATTTTTGCTGGGACCATCTACCTGCCTTATACTACACTTAGGAAAAAGTATTACATATG
GTTTATTTTGAAACTTCAAGTATTATTGCCTTAATGTCTCTTAACCCTGTTACACGCTGCTTGTAGACAT
GTTAATATAGTAATACCTTTATGATATATTGAGTTTAAGGACTACTCTTTTTCTGTTTTATCATGTATGC
ATTATTTTGTATATGTACAGGGCAAGTAGGTATATAATTTGATAAAGTTGCAATTGAAATATTATTAACA
GAAGATGTAAGAAATTTCTGCATGGTCTAAATCTTTGTGTACTTTATTTGTAAATTATTTGCCCTGGAGT
TTTAGAAAATAGTTTCTGAATTTTAAACTTGCTGGATTCATGCAGCCAGCTTTGCAGGTTATCAGAGATC
AAAGATTGTAATAATAATTTTGTAAATTGTAAGCAAAAAGTTATTTTTATATTATATACAGTCTAATTGT
TCATCCTAATTGTTCCTGTTTTCATCTAGTCAGAGATTCAGTAAGTGCCTTGGAACAATATTGAATTCTC
TTAGCTTGTGTGTGTTTCTTTAATATTTGAACTCAAGTGGATTAGAAGACTATCAAAATACATGTATGT
TTCAGGATATTTGACCTGTCATTAAAAAAAACAAACAGTTTTACAGTGCCAAAAAAAAAAAAAAA

Human PTPN12 mRNA sequence - var3 (public gi: 18375651) (SEQ ID NO: 125)
AGCGACCGCAGCCGGGGGGACGCGGGAGGATGGAGCAAGTGGAGATCCTGAGGAAATTCATCCAGAGGGT
CCAGGCCATGAAGAGTCCTGACCACAATGGGGAGGACAACTTCGCCCGGGACTTCATGCGGTTAAGAAGA
TTGTCTACCAAATATAGAACAGAAAAGATATATCCCACAGCCACTGGAGAAAAAGAAGAAAATGTTAAAA
AGAACAGATACAAGGACATACTGCCATTTGATCACAGCCGAGTTAAATTGACATTAAAGACTCCTTCACA
AGATTCAGACTATATCAATGCAAATTTTATAAAGGGCGTCTATGGGCCAAAAGCATATGTAGCAACTCAA
GGACCTTTAGCAAATACAGTAATAGATTTTTGGAGGATGATATGGGAGTATAATGTTGTGATCATTGTAA
TGGCCTGCCGAGAATTTGAGATGGGAAGGAAAAAATGTGAGCGCTATTGGCCTTTGTATGGAGAAGACCC
CATAACGTTTGCACCATTTAAAATTTCTTGTGAGGATGAACAAGCAAGAACAGACTACTTCATCAGGACA
CTCTTACTTGAATTTCAAAATGAATCTCGTAGGCTGTATCAGTTTCATTATGTGAACTGGCCAGACCATG
ATGTTCCTTCATCATTTGATTCTATTCTGGACATGATAAGCTTAATGAGGAAATATCAAGAACATGAAGA
TGTTCCTATTTGTATTCATTGCAGTGCAGGCTGTGGAAGAACAGGTGCCATTTGTGCCATAGATTATACG
TGGAATTTACTAAAAGCTGGGAAAATACCAGAGGAATTTAATGTATTTAATTTAATACAAGAAATGAGAA
CACAAAGGCATTCTGCAGTACAAACAAAGGAGCAATATGAACTTGTTCATAGAGCTATTGCCCAACTGTT
TGAAAAACAGCTACAACTATATGAAATTCATGGAGCTCAGAAAATTGCTGATGGAGTGAATGAAATTAAC
ACTGAAAACATGATCAGCTCCATAGAGCCTGAAAAACAAGATTCTCCTCCTCCAAAACCACCAAGGACCC
GCAGTTGCCTTGTTGAAGGGGATGCTAAAGAAGAAATACTGCAGCCACCGGAACCTCATCCAGTGCCACC
CATCTTGACACCTTCTCCCCCTTCAGCTTTTCCAACAGTCACTACTGTGTGGCAGGACAATGATAGATAC
CATCCAAAGCCAGTGTTGCATATGGTTTCATCAGAACAACATTCAGCAGACCTCAACAGAAACTATAGTA
AATCAACAGAACTTCCAGGGAAAAATGAATCAACAATTGAACAGATAGATAAAAAATTGGAACGAAATTT
AAGTTTTGAGATTAAGAAGGTCCCTCTCCAAGAGGGACCAAAAAGTTTTGATGGGAACACACTTTTGAAT
AGGGGACATGCAATTAAAATTAAATCTGCTTCACCTTGTATAGCTGATAAAATCTCTAAGCCACAGGAAT
TAAGTTCAGATCTAAATGTCGGTGATACTTCCCAGAATTCTTGTGTGGACTGCAGTGTAACACAATCAAA
CAAAGTTTCAGTTACTCCACCAGAAGAATCCCAGAATTCAGACACACCTCCAAGGCCAGACCGCTTGCCT
CTTGATGAGAAAGGACATGTAACGTGGTCATTTCATGGACCTGAAAATGCCATACCCATACCTGATTTAT
CTGAAGGCAATTCCTCAGATATCAACTATCAAACTAGGAAAACTGTGAGTTTAACACCAAGTCCTACAAC

ACAAGTTGAAACACCTGATCTTGTGGATCATGATAACACTTCACCACTCTTCAGAACACCCCTCAGTTTT
ACTAATCCACTTCACTCTGATGACTCAGACTCAGATGAAAGAAACTCTGATGGTGCTGTGACCCAGAATA
AAACTAATATTTCAACAGCAAGTGCCACAGTTTCTGCTGCCACTAGTACTGAAAGCATTTCTACTAGGAA
AGTATTGCCAATGTCCATTGCTAGACATAATATAGCAGGAACAACACATTCAGGTGCTGAAAAAGATGTT
GATGTTAGTGAAGATTCACCTCCTCCCCTACCTGAAAGAACTCCTGAATCGTTTGTGTTAGCAAGTGAAC
ATAATACACCTGTAAGATCGGAATGGAGTGAACTTCAAAGTCAGGAACGATCTGAACAAAAAAAGTCTGA
AGGCTTGATAACCTCTGAAAATGAGAAATGTGATCATCCAGCGGGAGGTATTCACTATGAAATGTGCATA
GAATGTCCACCTACTTTCAGTGACAAGAGAGAACAAATATCAGAAAATCCAACAGAAGCCACAGATATTG
GTTTTGGTAATCGATGTGGAAAACCCAAAGGACCAAGAGATCCACCTTCAGAATGGACATGATTCAGGGA
GCTAGAAGACACTTTAAGTTATACTGGAAAATTCAGGTGCCACTGAAAGCCAGATTTATAGTATTCCATC
TTTAATATGTGGGACTAACAGCAGTGTAGATTGTTACCTTAATATTTTTGCTGGGACCATCTACCTGCC
TTATACTACACTTAGGAAAAAGTATTACATATGGTTTATTTTGAAACTTCAAGTATTATTGCCTTAATGT
CTCTTAACCCTGTTACACGCTGCTTGTAGACATGTTAATATAGTAATACCTTTATGATATATTGAGTTTA
AGGACTACTCTTTTTCTGTTTTATCATGTATGCATTATTTTGTATATGTACAGGGCAAGTAGGTATATAA
TTTGATAAAGTTGCAATTGAAATATTATTAACAGAAGATGTAAGAAATTTCTGCATGGTCTAAATCTTTG
TGTACTTTATTTGTAAATTATTTGCCCTGGAGTTTTAGAAAATAGTTTCTGAATTTTAAACTTGCTGGAT
TCATGCAGCCAGCTTTGCAGGTTATCAGAGATCAAAGATTGTAATAATAATTTTGTAAATTGTAAGCAAA
AAGTTATTTTTATATTATATACAGTCTAATTGTTCATCCTAATTGTTCCTGTTTTCATCTAGTCAGAGAT
TCAGTAAGTGCCTTGGAACAATATTGAATTCTCTTAGCTTGTGTGTGTTTCTTTAATATTTGAACTCAAG
TGGGATTAGAAGACTATCAAAATACATGTATGTTTCAGGATATTTGACCTGTCATTAAAAAAAACAAACA
GTTTTACAGTG

Human PTPN12 mRNA sequence - var4 (public gi: 545651) (SEQ ID NO: 126)
GTTAAAAGGAACAGATACAAGGACATACTGCCATTTGATCACAGCCGAGTTAAATTGACATTAAAGACTC
CTTCACAAGATTCAGACTATATCAATGCAAATTTTATAAAGGGCGTCTATGGCCAAAAGCATATGTAGC
AACTCAAGGACCTTTAGCAAATACAGTAATAGATTTTTGGAGGATGGTATGGGAGTATAATGTTGTGATC
ATTGTAATGGCCTGCCGAGAATTTGA

Human PTPN12 mRNA sequence - var5 (public gi: 19683965) (SEQ ID NO: 127)
GGGACTTCACCACTCTTCAGAACACCCCTCAGTTTTAGTAATCCACTTCACTCTGATGACTCAGACTCAG
ATGAAAGAAACTCTGATGGTGCTGTGACCCAGAATAAAACTAATATTTCAACAGCAAGTGCCACAGTTTC
TGCTGCCACTAGTACTGAAAGCATTTCTACTAGGAAAGTATTGCCAATGTCCATTGCTAGACATAATATA
GCAGGAACAACACATTCAGGTGCTGAAAAAGATGTTGATGTTAGTGAAGATTCACCTCCTCCCCTACCTG
AAAGAACTCCTGAATCGTTTGTGTTAGCAAGTGAACATAATACACCTGTAAGATCGGAATGGAGTGAACT
TCAAAGTCAGGAACGATCTGAACAAAAAAAGTCTGAAGGCTTGATAACCTCTGAAAATGAGAAATGTGAT
CATCCAGCGGGAGGTATTCACTATGAAATGTGCATAGAATGTCCACCTACTTTCAGTGACAAGAGAGAAC
AAATATCAGAAAATCCAACAGAAGCCACAGATATTGGTTTTGGTAATCGATGTGGAAAACCCAAAGGACC
AAGAGATCCACCTTCAGAATGGACATGATTCAGGGAGCTAGAAGACACTTTAAGTTATACTGGAAAATTC
AGGTGCCACTGAAAGCCAGATTTATAGTATTCCATCTTTAATATGTGGGACTAACAGCAGTGTAGATTGT
TACCTTAATATTTTTGCTGGGACCATCTACCTGCCTTATACTACACTTAGGAAAAAGTATTACATATGG
TTTATTTTGAAACTTCAAGTATTATTGCCTTAATGTCTCTTAACCCTGTTACACGCTGCTTGTAGACATG
TTAATATAGTAATACCTTTATGATATATTGAGTTTAAGGACTACTCTTTTTCTGTTTTATCATGTATGCA
TTATTTTGTATATGTACAGGGCAAGTAGGTATATAATTTGATAAAGTTGCAATTGAAATATTATTAACAG
AAGATGTAAGAAATTTCTGCATGGTCTAAATCTTTGTGTACTTTATTTGTAAATTATTTGCCCTGGAGTT
TTAGAAAATAGTTTCTGAATTTTAAACTTGCTGGATTCATGCAGCCAGCTTTGCAGGTTATCAGAGATCA
AAGATTGTAATAATAATTTTGTAAATTGTAAGCAAAAAGTTATTTTTATATTATATACAGTCTAATTGTT
CATCCTAATTGTTCCTGTTTTCATCTAGTCAGAGATTCAGTAAGTGCCTTGGAACAATATTGAATTCTCT
TAGCTTGTGTGTGTTTCTTTAATATTTGAACTCAAGTGGGATTAGAAGACTATCAAAATACATGTATGTT
TCAGGATATTTGACCTGTCATTAAAAAAAACAAACAGTTTTACAATAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAA

Human PTPN12 mRNA sequence - var6 (public gi: 220033) (SEQ ID NO: 128)
GCCGGGGGGACGCGGGAGGATGGAGCAAGTGGAGATCCTGAGGAAATTCATCCAGAGGGTCCAGGCCATG
AAGAGTCCTGACCACAATGGGGAGGACAACTTCGCCCGGGACTTCATGCGGTTAAGAAGATTGTCTACCA
AATATAGAACAGAAAAGATATATCCCACAGCCACTGGAGAAAAAGAAGAAAATGTTAAAAAGAACAGATA
CAAGGACATACTGCCATTTGATCACAGCCGAGTTAAATTGACATTAAAGACTCCTTCACAAGATTCAGAC
TATATCAATGCAAATTTTATAAAGGGCGTCTATGGCCAAAAGCATATGTAGCAACTCAAGGACCTTTAG
CAAATACAGTAATAGATTTTTGGAGGATGGTATGGGAGTATAATGTTGTGATCATTGTAATGGCCTGCCG
AGAATTTGAGATGGGAAGGAAAAAATGTGAGCGCTATTGGCCTTTGTATGGAGAAGACCCCATAACGTTT
GCACCATTTAAAATTTCTTGTGAGGATGAACAAGCAAGAACAGACTACTTCATCAGGACACTCTTACTTG
AATTTCAAAATGAATCTCGTAGGCTGTATCAGTTTCATTATGTGAACTGGCCAGACCATGATGTTCCTTC
ATCATTTGATTCTATTCTGGACATGATAAGCTTAATGAGGAAATATCAAGAACATGAAGATGTTCCTATT

TGTATTCATTGCAGTGCAGGCTGTGGAAGAACAGGTGCCATTTGTGCCATAGATTATACGTGGAATTTAC
TAAAAGCTGGGAAAATACCAGAGGAATTTAATGTATTTAATTTAATACAAGAAATGAGAACACAAAGGCA
TTCTGCAGTACAAACAAAGGAGCAATATGAACTTGTTCATAGAGCTATTGCCCAACTGTTTGAAAAACAG
CTACAACTATATGAAATTCATGGAGCTCAGAAAATTGCTGATGGAGTGAATGAAATTAACACTGAAAACA
TGGTCAGCTCCATAGAGCCTGAAAAACAAGATTCTCCTCCTCCAAAACCACCAAGGACCCGCAGTTGCCT
TGTTGAAGGGGATGCTAAAGAAGAAATACTGCAGCCACCGGAACCTCATCCAGTGCCACCCATCTTGACA
CCTTCTCCCCCTTCAGCTTTTCCAACAGTCACTACTGTGTGGCAGGACAATGATAGATACCATCCAAAGC
CAGTGTTGCATATGGTTTCATCAGAACAACATTCAGCAGACCTCAACAGAAACTATAGTAAATCAACAGA
ACTTCCAGGGAAAAATGAATCAACAATTGAACAGATAGATAAAAAATTGGAACGAAATTTAAGTTTTGAG
ATTAAGAAGGTCCCTCTCCAAGAGGGACCAAAAAGTTTTGATGGGAACACACTTTTGAATAGGGGACATG
CAATTAAAATTAAATCTGCTTCACCTTGTATAGCTGATAAAATCTCTAAGCCACAGGAATTAAGTTCAGA
TCTAAATGTCGGTGATACTTCCCAGAATTCTTGTGTGGACTGCAGTGTAACACAATCAAACAAAGTTTCA
GTTACTCCACCAGAAGAATCCCAGAATTCAGACACACCTCCAAGGCCAGACCGCTTGCCTCTTGATGAGA
AAGGACATGTAACGTGGTCATTTCATGGACCTGAAAATGCCATACCCATACCTGATTTATCTGAAGGCAA
TTCCTCAGATATCAACTATCAAACTAGGAAAACTGTGAGTTTAACACCAAGTCCTACAACACAAGTTGAA
ACACCTGATCTTGTGGATCATGATAACACTTCACCACTCTTCAGAACACCCCTCAGTTTTACTAATCCAC
TTCACTCTGATGACTCAGACTCAGATGAAAGAAACTCTGATGGTGCTGTGACCCAGAATAAAACTAATAT
TTCAACAGCAAGTGCCACAGTTTCTGCTGCCACTAGTACTGAAAGCATTTCTACTAGGAAAGTATTGCCA
ATGTCCATTGCTAGACATAATATAGCAGGAACAACACATTCAGGTGCTGAAAAAGATGTTGATGTTAGTG
AAGATTCACCTCCTCCCCTACCTGAAAGAACTCCTGAATCGTTTGTGTTAGCAAGTGAACATAATACACC
TGTAAGATCGGAATGGAGTGAACTTCAAAGTCAGGAACGATCTGAACAAAAAAAGTCTGAAGGCTTGATA
ACCTCTGAAAATGAGAAATGTGATCATCCAGCGGGAGGTATTCACTATGAAATGTGCATAGAATGTCCAC
CTACTTTCAGTGACAAGAGAGAACAAATATCAGAAAATCCAACAGAAGCCACAGATATTGGTTTTGGTAA
TCGATGTGGAAAACCCAAAGGACCAAGAGATCCACCTTCAGAATGGACATGATTCAGGGAGCTAGAAGAC
ACTTTAAGTTATACTGGAAAATTCAGGTGCCACTGAAAGCCAGATTTATAGTATTCCATCTTTAATATGT
GGGACTAACAGCAGTGTAGATTGTTACCTTAATATTTTTGCTGGGACCATCTACCTGCCTTATACTACA
CTTAGGAAAAAGTATTACATATGGTTTATTTGAAACTTCAAGTATTATTGCCTTAATGTCTCTTAACCC
TGTTACACGCTGCTTGTAGACATGTTAATATAGTAATACCTTTATGATATATTGAGTTTAAGGACTACCC
TTTTTCTGTTTTATCATGTATTCATTATTTGTATATGTACAGGGCAAGTAGGTATATAATTTGATAAAG
TTGCAATTGAAATATTATTAACAGAAGATGTAAGAAATTTCTGCATGGTCTAAATCTTTGTGTACTTTAT
TTGTAAATTATTTGCCCTGGAGTTTTAGAAAATAGTTTCTGAATTTTAAACTTGCTGGATTCATGCAGCC
AGCTTTGCAGGTTATCAGAGATCAAAGATTGTAATAATAATTTGTAAATTGTAAGCAAACATTCTGC

Human PTPN12 protein sequence - var1 (public gi: 220034) (SEQ ID NO: 266)
MEQVEILRKFIQRVQAMKSPDHNGEDNFARDFMRLRRLSTKYRTEKIYPTATGEKEENVKKNRYKDILPF
DHSRVKLTLKTPSQDSDYINANFIKGVYGPKAYVATQGPLANTVIDFWRMVWEYNVVIIVMACREFEMGR
KKCERYWPLYGEDPITFAPFKISCEDEQARTDYFIRTLLLEFQNESRRLYQFHYVNWPDHDVPSSFDSIL
DMISLMRKYQEHEDVPICIHCSAGCGRTGAICAIDYTWNLLKAGKIPEEFNVFNLIQEMRTQRHSAVQTK
EQYELVHRAIAQLFEKQLQLYEIHGAQKIADGVNEINTENMVSSIEPEKQDSPPPKPPRTRSCLVEGDAK
EEILQPPEPHPVPPILTPSPPSAFPTVTTVWQDNDRYHPKPVLHMVSSEQHSADLNRNYSKSTELPGKNE
STIEQIDKKLERNLSFEIKKVPLQEGPKSFDGNTLLNRGHAIKIKSASPCIADKISKPQELSSDLNVGDT
SQNSCVDCSVTQSNKVSVTPPEESQNSDTPPRPDRLPLDEKGHVTWSFHGPENAIPIPDLSEGNSSDINY
QTRKTVSLTPSPTTQVETPDLVDHDNTSPLFRTPLSFTNPLHSDDSDSDERNSDGAVTQNKTNISTASAT
VSAATSTESISTRKVLPMSIARHNIAGTTHSGAEKDVDVSEDSPPPLPERTPESFVLASEHNTPVRSEWS
ELQSQERSEQKKSEGLITSENEKCDHPAGGIHYEMCIECPPTFSDKREQISENPTEATDIGFGNRCGKPK
GPRDPPSEWT

Human PTPN12 protein sequence - var2 (public gi: 7689910) (SEQ ID NO: 267)
VKRNRYKDILPFDHSRVKLTLKTPSQDSDYINANFIKGVYGPKAYVATQGPLANTVIDFWRMVWEYNVVI
IVMACREF

Human PTPN12 protein sequence - var3 (public gi: 292409) (SEQ ID NO: 268)
MEQVEILRKFIQRVQAMKSPDHNGEDNFARDFMRLRRLSTKYRTEKIYPTATGEKEENVKKNRYKDILPF
DHSRVKLTLKTPSQDSDYINANFIKGVYGPKAYVATQGPLANTVIDFWRMIWEYNVVIIVMACREFEMGR
KKCERYWPLYGEDPITFAPFKISCEDEQARTDYFIRTLLLEFQNESRRLYQFHYVNWPDHDVPSSFDSIL
DMISLMRKYQEHEDVPICIHCSAGCGRTGAICAIDYTWNLLKAGKIPEEFNVFNLIQEMRTQRHSAVQTK
EQYELVHRAIAQLFEKQLQLYEIHGAQKIADGVNEINTENMISSIEPEKQDSPPPKPPRTRSCLVEGDAK
EEILQPPEPHPVPPILTPSPPSAFPTVTTVWQDNDRYHPKPVLHMVSSEQHSADLNRNYSKSTELPGKNE
STIEQIDKKLERNLSFEIKKVPLQEGPKSFDGNTLLNRGHAIKIKSASPCIADKISKPQELSSDLNVGDT
SQNSCVDCSVTQSNKVSVTPPEESQNSDTPPRPDRLPLDEKGHVTWSFHGPENAIPIPDLSEGNSSDINY
QTRKTVSLTPSPTTQVETPDLVDHDNTSPLFRTPLSFTNPLHSDDSDSDERNSDGAVTQNKTNISTASAT
VSAATSTESISTRKVLPMSIARHNIAGTTHSGAEKDVDVSEDSPPPLPERTPESFVLASEHNTPVRSEWS

ELQSQERSEQKKSEGLITSENEKCDHPAGGIHYEMCIECPPTFSDKREQISENPTEATDIGFGNRCGKPK
GPRDPPSEWT

Human PTPN12 pray sequence - var1 (SEQ ID NO: 129)
GTTTGGNATNCTACAGGGNATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATTCGATGAT
GAAGATACCCCACCAAACCCAAAAAAAGAGATCTTTAATACGACTCACTATAGGGCGAGCGCCGCCATGG
AGTACCCATACGACGTACCAGATTACGCTCATATGGCCATGGAGGCCAGTGAATTCCACCCAAGCAGTGG
TATCAACGCAGAGTGGAATTATGGGCGGCGGCGGCGGCGTCTCCGACGGAGGAGGAGGGCGGGGAAGGAG
GATGGAGCAAACAAAGTTTCAGTTACTCCACCAGAAGAATCCCAGAATTCAGACACACCTCCAAGGCCAG
ACCGCTTGCCTCTTGATGAGAAAGGACATGTAACGTGGTCATTTCATGGACCTGAAAATGCCATACCCAT
ACCTGATTTATCTGAAGGCAATTCCTCAGATATCAACTATCAAACTAGGAAAACTGTGAGTTTAACACCA
AGTCCTACAACACAAGTTGAANGCACCTGATCTTGTGGATCATGATAACGCTTCACCACTCTTCAGAACA
CCCCTCANTTTTACTAATCCACTTCACTCTNATGACTCANACTCANATGAAAGAAACTCTGATGGTGCTG
TGACCCANAATAAAACTAATATTTCAACAGCAAGTGCCACAGTTTCTGCTGCCACTANTACTGAAAGCAT
TTCTACTAGGAAAGTATTGCCNATGTCCATTGCTAGACNTANTATANCAGGAACANACATTCAGGTGCTG
AAAAAAANTTNATGTTNNTGAANATTNCCTNCTCCCNNCCTNAANAACTCC

<u>Unigene Name</u>: RALA <u>Unigene ID</u>: Hs.6906 <u>Clone ID</u>: 3GD_1106

Human RALA mRNA sequence - var1 (public gi: 35845) (SEQ ID NO: 130)
ATGGCTGCAAATAAGCCCAAGGGTCAGAATTCTTTGGCTTTACACAAAGTCATCATGGTGGGCAGTGGTG
GCGTGGGCAAGTCAGCTCTGACTCTACAGTTCATGTACGATGAGTTTGTGGAGGACTATGAGCCTACCAA
AGCAGACAGCTATCGGAAGAAGGTAGTGCTAGATGGGGAGGAAGTCCAGATCGATATCTTAGATACAGCT
GGGCAGGAGGACTACGCTGCAATTAGAGACAACTACTTCCGAAGTGGGGAGGGGTTCCTCTGTGTTTTCT
CTATTACAGAAATGGAATCCTTTGCAGCTACAGCTGACTTCAGGGAGCAGATTTTAAGAGTAAAAGAAGA
TGAGAATGTTCCATTTCTACTGGTTGGTAACAAATCAGATTTAGAAGATAAAAGACAGGTTTCTGTAGAA
GAGGCAAAAAACAGAGCTGAGCAGTGGAATGTTAACTACGTGGAAACATCTGCTAAAACACGAGCTAATG
TTGACAAGGTATTTTTTGATTTAATGAGAGAAATTCGAGCGAGAAAGATGGAAGACAGCAAAGAAAAGAA
TGGAAAAAAGAAGAGGAAAAGTTTAGCCAAGAGAATCAGAGAAAGATGCTGCATTTTATAA

Human RALA mRNA sequence - var2 (public gi: 24980846) (SEQ ID NO: 131)
CCGCTCCCCAGAGCAAAGCGTCGGAGTCCTCCTCCTCCTTCTCCTCCTCCTCCTCCTCCTCCTCCAGCCG
CCCAGGCTCCCCCGCCACCCGTCAGACTCCTCCTTCGACCGCTCCCGGCGCGGGGCCTTCCAGGCGACAA
GGACCGAGTACCCTCCGGCCGGAGCCACGCAGCCGCGGCTTCCGGAGCCCTCGGGGCGGCGGACTGGCTC
GCGGTGCAGATTCTTCTTAATCCTTTGGTGAAAACTGAGACACAAAATGGCTGCAAATAAGCCCAAGGGT
CAGAATTCTTTGGCTTTACACAAAGTCATCATGGTGGGCAGTGGTGGCGTGGGCAAGTCAGCTCTGACTC
TACAGTTCATGTACGATGAGTTTGTGGAGGACTATGAGCCTACCAAAGCAGACAGCTATCGGAAGAAGGT
AGTGCTAGATGGGGAGGAAGTCCAGATCGATATCTTAGATACAGCTGGGCAGGAGGACTACGCTGCAATT
AGAGACAACTACTTCCGAAGTGGGGAGGGGTTCCTCTGTGTTTTCTCTATTACAGAAATGGAATCCTTTG
CAGCTACAGCTGACTTCAGGGAGCAGATTTTAAGAGTAAAAGAAGATGAGAATGTTCCATTTCTACTGGT
TGGTAACAAATCAGATTTAGAAGATAAAAGACAGGTTTCTGTAGAAGAGGCAAAAAACAGAGCTGAGCAG
TGGAATGTTAACTACGTGGAAACATCTGCTAAAACACGAGCTAATGTTGACAAGGTATTTTTTGATTTAA
TGAGAGAAATTCGAGCGAGAAAGATGGAAGACAGCAAAGAAAAGAATGGAAAAAAGAAGAGGAAAAGTTT
AGCCAAGAGAATCAGAGAAAGATGCTGCATTTTATAATCAAAGCCCAAACTCCTTTCTTATCTTGACCAT
ACTAATAAATATAATTTATAAGCATTGCCATTGAAGGCTTAATTGACTGAAATTACTTTAACATTTTGGA
AATTGTTGTATATCACTAAAAGCATGAATTGGAACTGCAATGAAAGTCAAATTTACTTTAAAAAGAAATT
AATATGGCTTCACCAAGAAGCAAAGTTCAACTTATTTCATAATTGCCTACATTTATCATGGTCCTGAATG
TAGCGTGTAAGCTTGTGTTTCTTGGGCAGTCTTTCTTGAAATTGAAGAGGTGAAATGGGGGTGGGGAGTG
GGAGGAAAGGTGACTTCCTCTGGTGTTTATTATAAAGCTTAAATTTTATATCATTTTAAAATGTCTTGGT
CTTCTACTGCCTTGAAAAATGACAATTGTGAACATGATAGTTAAACTACCACTTTTTTTAACCATTATTA
TGCAAAATTTAGAAGAAAAGTTATTGGCATGGTTGTTGCATATAGTTAAACTGAGAGTAATTCATCTGTG
AATCTGCTTTAATTACCTGGTGAGTAACTTAGAAAAGTGGTGTAAACTTGTACATGGAATTTTTTGAATA
TGCCTTAATTTAGAAACTGAAAAATATCTGGTTATATCATTCTGGGTGTGTTCTTACTGACACCAGGGGT
CCGCTGCCCCATGTGTCCTGGTGAGAAAATATATGCCTGGCACAGCTTTTGTATAGAAAATTCTTGAGAA
GTAACTGTCCGCTAGAAGTCTGTCCAAATTTAAAATGTGTGCCATATTCTGGTTCTTGAAAATAAGATTC
CAGAGCTCTTTGATCGCTTTTAATAAACTGCAAGTTCATTTTAAATGAAGGGCCAGCATATATACTTGCA
AGATAATTTTCAGCTGCAAGGATTCAGCACCAGTTATGTTTGAATGAACCCTCCTTTTCTCTGAGATTCT
GGTCCCTGGAAATCCCTTTCTGCTAGTGGTGAGCATGTAAGTGTTAAGTTTTTAATCTGGGAGCAGGGCA
TAGGAAGAAAATGTCAGTAGTGCTAATGCATTTTGCACTAGAACGCTTCGGGAAAATATTCATGCTTGCC
ATCTGTTCATTTCTAAATTTATATTCATAAAGTTACAGTTTGATACAGGAATTATTAGGAGTAATTCTTT
TCTGTTTCTGTTTATAATGAAGAACACTGTAGCTACATTTTCAGAAGTTAACATCAAGCCATCAAACCTG
GGTATAGTGCAGAAAACGTGGCACACACTGACCACACATTAGGCTGTGTCACCATTGTGTGGTGTACCTG

CTGGAAGAATTCTAGCATGCTACTTGGGGACATAATTTCAGTGGAAATATGCCACTGACCGATTTTTTT
TTTTTCCTCTTTGCAGTGGGGCTAGGACAGTTGATTCAACAAAGTATTTTTTTCTTTTTTCTCAGTCCTA
ATTTGAACAGGTCAAAGATGTGTTCAGGCATTCCAGGTAACAGGTGTGTATGTAAAGTTAAAAATAGGCT
TTTTAGGAACTCACTCTTTAGATATTTACATCCAGCTTCTCATGTTAAATATTTGTCCTTAAAGGGTTTG
AGATGTACATCTTTCATTTCGTATTTCTCATAGGCTATGCCATGTGCGGAATTCAAGTTACCAATGTAAC
ACTGGCCAGCGGGCCCAGCAATCTCCATGTGTACTTATTACAGTCTTATTTAACCAGGGGTCCTAACCAC
TAACATTGTGACTTTGCTTTGAGACCTTTCCTCTCCTGGGTACTGAGGTGCTATGAAGCCAACTGACAAA
GATGCATCACGTGTCTTAGGCTGATGCCACTACCCGATTTGTTTATTTGCAATTTGAGCCATTTAAAGAC
CAATAAACTTCCTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAA

Human RALA mRNA sequence - var3 (public gi: 3483427) (SEQ ID NO: 132)
ATAATCAAAGCCCAAACTCCTTTCTTATCTTGACCATACTAATAAATATAATTTATAAGCATTGCCATTG
AAGGCTTAATTGACTGAAATTACTTTAACATTTTGGAAATTGTTGTATATCACTAAAAGCATGAATTGGA
ACTGCAATGAAAGTCAAATTTACTTTAAAAAGAAATTAATATGGCTTCACCAAGAAGCAAAGTTCAACTT
ATTTCATAATTGCCTACATTTATCATGGTCCTGAATGTAGCGTGTAAGCTTGTGTTTCTTGGGCAGTCTT
TCTTGAAATTGAAGAGGTGAAATGGGGGTGGGGAGTGGGAGGAAAGGTGACTTCCTCTGGTGTTTATTAT
AAAGCTTAAATTTTATATCATTTTAAAATGTCTTGGTCTTCTACTGCCTTGAAAAATGACAATTGTGAAC
ATGATAGTTAAACTACCACTTTTTTTAACCATTATTATGCAAAAAAAAAA

Human RALA mRNA sequence - var4 (public gi: 20147712) (SEQ ID NO: 133)
ATGGTCGACTACCTAGCAAATAAGCCCAAGGGTCAGAATTCTTTGGCTTTACACAAAGTCATCATGGTGG
GCAGTGGTGGCGTGGGCAAGTCAGCTCTGACTCTACAGTTCATGTACGATGAGTTTGTGGAGGACTATGA
GCCTACCAAAGCAGACAGCTATCGGAAGAAGGTAGTGCTAGATGGGGAGGAAGTCCAGATCGATATCTTA
GATACAGCTGGGCAGGAGGACTACGCTGCAATTAGAGACAACTACTTCCGAAGTGGGGAGGGGTTCCTCT
GTGTTTTCTCTATTACAGAAATGGAATCCTTTGCAGCTACAGCTGACTTCAGGGAGCAGATTTTAAGAGT
AAAAGAAGATGAGAATGTTCCATTTCTACTGGTTGGTAACAAATCAGATTTAGAAGATAAAAGACAGGTT
TCTGTAGAAGAGGCAAAAAACAGAGCTGAGCAGTGGAATGTTAACTACGTGGAAACATCTGCTAAAACAC
GAGCTAATGTTGACAAGGTATTTTTTGATTTAATGAGAGAAATTCGAGCGAGAAAGATGGAAGACAGCAA
AGAAAAGAATGGAAAAAAGAAGAGGAAAAGTTTAGCCAAGAGAATCAGAGAAAGATGCTGCATTTTATAA

Human RALA mRNA sequence - var5 (public gi: 10439805) (SEQ ID NO: 134)
AGAATGGAAAAAAGAAGAGGAAAAGTTTAGCCAAGAGAATCAGAGAAAGATGCTGCATTTTATAATCAAA
GCCCAAACTCCTTTCTTATCTTGACCATACTAATAAATATAATTTATAAGCATTGCCATTGAAGGCTTAA
TTGACTGAAATTACTTTAACATTTTGGAAATTGTTGTATATCACTAAAAGCATGAATTGGAACTGCAATG
AAAGTCAAATTTACTTTAAAAAGAAATTAATATGGCTTCACCAAGAAGCAAAGTTCAACTTATTTCATAA
TTGCCTACATTTATCATGGTCCTGAATGTAGCGTGTAAGCTTGTGTTTCTTGGGCAGTCTTTCTTGAAAT
TGAAGAGGTGAAATGGGGGTGGGGAGTGGGAGGAAAGGTGACTTCCTCTGGTGTTTATTATAAAGCTTAA
ATTTTATATCATTTTAAAATGTCTTGGTCTTCTACTGCCTTGAAAAATGACAATTGTGAACATGATAGTT
AAACTACCACTTTTTTTAACCATTATTATGCAAAATTTAGAAGAAAAGTTATTGGCATGGTTGTTGCATA
TAGTTAAACTGAGAGTAATTCATCTGTGAATCTGCTTTAATTACCTGGTGAGTAACTTAGAAAAGTGGTG
TAAACTTGTACATGGAATTTTTTGAATATGCCTTAATTTAGAAACTGAAAAATATCCGGTTATATCATTC
TGGGTGTGTTCTTACTGACACCAGGGGTCCGCTGCCCCATGTGTCCTGGTGAGAAAATATATGCCTGGCA
CAGCTTTTGTATAGAAAATTCTTGAGAAGTAACTGTCCGCTAGAAGTCTGTCCAAATTTAAAATGTGTGC
CATATTCTGGTTCTTGAAAATAAGATTCCAGAGCTCTTTGATCGCTTTTAATAAACTGCAAGTTCATTTT
AATTGAAGGGCCAGCATATATACTTGCAAGATAATTTTCAGCTGCAAGGATTCAGCACCAGTTATGTTTG
AATGAACCCTCCTTTTCTCTGAGATTCTGGTCCCTGGAAATCCCTTTCTGCTAGTGGTGAGCATGTAAGT
GTTAAGTTTTTAATCTGGGAGCAGGGCATAGGAAGAAAATGTCAGTAGTGCTAATGCATTTTGCACTAGA
ACGCTTCGGGAAAATATTCATGCTTGCCATCTGTTCATTTCTAAATTTATATTCATAAAGTTACAGTTTG
ATACAGGAATTATTAGGAGTAATTCTTTTCTGTTTCTGTTTATAATGAAGAACACTGTAGCTACATTTTC
AGAAGTTAACATCAAGCCATCAAACCTGGGTATAGTGCAGAAGACGTGGCACACACTGACCACACATTAG
GCTGTGTCACCATTGTGTGGTGTACCTGCTGGAAGAATTCTAGCATGCTACTTGGGGACATAATTTCAGT
GGGAAATATGCCACTGACCGATTTTTTTTTTTTTCCTCTTTGCAGTGGGGCTAGGACAGTTGATTCAACA
AAGTATTTTTTTCTTTTTTCTCAGTCCTAATTTGGACAGGTCAAAGATGTGTTCAGGCATTCCAGGTAAC
AGGTGTGTATGTAAAGTTAAAAATAGGCTTTTTAGGAACTCACTCTTTAGATATTTACATCCAGCTTCTC
ATGTTAAATATTTGTCCTTAAAGGGTTTGAGATGTACATCTTTCATTTCGTATTTCTCATAGGCTATGCC
ATGTGCGGAATTCAAGTTACCAATGTAACACTGGCCAGCGGGCCCAGCAATCTCCATGTGTACTTATTAC
AGTCTTATTTAACCAGGGGTCCTAACCACTAACATTGTGACTTTGCTTTGAGACCTTTCCTCTCCTGGGT
ACTGAGGTGCTATGAAGCCAACTGACAAAGATGCATCACGTGTCTTAGGCTGATGCCACTACCCGATTTG
TTTATTTGCAATTTGAGCCATTTAAAGACCAATAAACTTCCTTTTTTAAAAAAAAAAAAAAAAAAAAAA
AAAAA

Human RALA Protein sequence - var1 (public gi: 35846) (SEQ ID NO: 269)

MAANKPKGQNSLALHKVIMVGSGGVGKSALTLQFMYDEFVEDYEPTKADSYRKKVVLDGEEVQIDILDTA
GQEDYAAIRDNYFRSGEGFLCVFSITEMESFAATADFREQILRVKEDENVPFLLVGNKSDLEDKRQVSVE
EAKNRAEQWNVNYVETSAKTRANVDKVFFDLMREIRARKMEDSKEKNGKKKRKSLAKRIRERCCIL

Human RALA Protein sequence - var2 (public gi: 20147713) (SEQ ID NO: 270)
MVDYLANKPKGQNSLALHKVIMVGSGGVGKSALTLQFMYDEFVEDYEPTKADSYRKKVVLDGEEVQIDIL
DTAGQEDYAAIRDNYFRSGEGFLCVFSITEMESFAATADFREQILRVKEDENVPFLLVGNKSDLEDKRQV
SVEEAKNRAEQWNVNYVETSAKTRANVDKVFFDLMREIRARKMEDSKEKNGKKKRKSLAKRIRERCCIL

Unigene Name: SIAH1 **Unigene ID**: Hs.295923 **Clone ID**: 3GD_150

Human SIAH1 mRNA sequence - var1 (public gi: 27503513) (SEQ ID NO: 135)

CCAGCGCGTCGCCCCCTGCATCCGTGGCCTCCACTGGAGCTGGGCAGGACCCTACCCAGTGAATCTGGAG
AAAACAAACTTGGAGACAGACGAAAGCTTAGGGCACATTGGAGGACAGCGCAGCTGTGGCTCCCATTTT
TGGAGATGCAGTCGAATTTGAGCTCACAGGGAGGTGTGGTTGCCTCCTGGGGATGGAAAGGCTTCCTTTC
TCCACCTCTGTAACTGGTGCTTCTGAGAAGTAAATGGTATTTGGATCCTGACCTCAGACGTGAATTTGGG
TCTTCTGTGCTTAGGAGCAGAAAGAGCCCAGGAGGGGCCTGTTCCTTTACTTCTTGGGGGAAACGCAATG
CGTGGCCTGACTTCTCATGACGGGAAAGGCTACTCCACCTTCTCTGTACTCCTGGAGGGGAGTCTTGTTC
ACATGTTTACCAGCGGCCAGGACAAGGAAGAGAAAAGAAATGAGCCGTCAGACTGCTACAGCATTACCTA
CCGGTACCTCGAAGTGTCCACCATCCCAGAGGGTGCCTGCCCTGACTGGCACAACTGCATCCAACAATGA
CTTGGCGAGTCTTTTTGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCCATTCTTCAATGTCAGAGT
GGCCATCTTGTTTGTAGCAACTGTCGCCCAAAGCTCACATGTTGTCCAACTTGCCGGGGCCCTTTGGGAT
CCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATTCAGTACTTTTCCCCTGTAAATATGCGTCTTCTGG
ATGTGAAATAACTCTGCCACACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCCTTATTCC
TGTCCGTGCCCTGGTGCTTCCTGTAAATGGCAAGGCTCTCTGGATGCTGTAATGCCCCATCTGATGCATC
AGCATAAGTCCATTACAACCCTACAGGGAGAGGATATAGTTTTTCTTGCTACAGACATTAATCTTCCTGG
TGCTGTTGACTGGGTGATGATGCAGTCCTGTTTTGGCTTTCACTTCATGTTAGTCTTAGAGAAACAGGAA
AAATACGATGGTCACCAGCAGTTCTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGAAAATT
TTGCTTACCGACTTGAGCTAAATGGTCATAGGCGACGATTGACTTGGGAAGCGACTCCTCGATCTATTCA
TGAAGGAATTGCAACAGCCATTATGAATAGCGACTGTCTAGTCTTTGACACCAGCATTGCACAGCTTTTT
GCAGAAAATGGCAATTTAGGCATCAATGTAACTATTTCCATGTGTTGAAATGGCAATCAAACATTTTCTG
GCCAGTGTTTAAAACTTCAGTTTCACAGAAAATAAGGCACCCATCTGTCTGCCAACCTAAAACTCTTTCG
GTAGGTGGAAGCTAGACACATGAAGGTAAATAAAAAGAAAGGCTGTTAAATACAGGAAACAGTTGCATGT
AGTAACACTAATATATTTAAAAATAAGTCAACAGTAAACCACTGAAAAAATATATGTATATACACCCAAG
ATGGGCATCTTTTGTATTAAGAAAGGAAGCATTGTAAAATAATTCTGAGTTTTGTGTTTGTTGTAGATTG
ATTGTATTGTTGAAAAAGTTTGTTTTTGCGTGGGAGTGTGTGCCTGCGTGGGTGTGTGCGTGTTTGGGTT
TTTTTCCTTTAACTGACAAGCCATCTTGAGTGGTCATGGGCCACTGCTTTTCCCTTTGTGAGTCAATACA
TAGTGCTGCTGTGTGCTTTTTTTGTGTGTATTTGCTAATTTTTATTAATTTAGTTTTTCATTAAATAAA
TTTGACTTTTCTGTAATTCAGGTTTTTCCTTTTTTTGTACCATTTTAAAGTTAGTATCTTTTGATATGCA
TATTTGTTTATGGTAAAAAATTTATAACGTGTTCAATATTTTCTTTTCCCCCATTAATCAGTTCATTAGA
AATATTTTAAAATCAGCTATTTTGTGAAGCCATGAGTTCCAGAAAGTAAAGGTGACATCGGAAAAAATAAT
CAAAAGCTATTTAAAGCATCTATAAGGTGCTCTCTTTCTGTCTTCTACAGATGAGTCACACCTTTGAGCT
TAATCTTTGAAAGGTTAGAGAATAAATTGATTTTTATAAATACTGCAAATCAGGCTTTTGTTTCCTTTTT
CAGATATCTTGGACAAATCACATATTTTAAAATTTGTTCTTGTATTTATTGGTTTTGCAGAAGAAGGCAT
CGTCATGCACAGTATTTGTAATTAAAAGCAAATCATTTGTTTAAAAAGGCAGTTTGCAAAAAATGTTTTT
GGTCTTTTATAATTCTCATTAAAAGAATATCTGTCAAATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAA

Human SIAH1 mRNA sequence - var2 (public gi: 4506946) (SEQ ID NO: 136)

GCGGCGGCCAGGGGGGAGCCGGGGCGGCCGTTGCGGGGCGCGCTCTCGAGAGGCGGCGGCGGCCAGGGTG
TCCCGTCGGTCTCGGCGCCGGGAAGAGGCGGTGGCGCTGCCCGCGGTGGCGGGGGTTGGCGACGGAGCGC
GTTGGTGCCAGGACCGGGGTCCGAGGCGCGCTCTCCGCCCACAGAAATGAGCCGTCAGACTGCTACAGCA
TTACCTACCGGTACCTCGAAGTGTCCACCATCCCAGAGGGTGCCTGCCCTGACTGGCACAACTGCATCCA
ACAATGACTTGGCGAGTCTTTTTGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCCATTCTTCAATG
TCAGAGTGGCCATCTTGTTTGTAGCAACTGTCGCCCAAAGCTCACATGTTGTCCAACTTGCCGGGGCCCT
TTGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATTCAGTACTTTTCCCCTGTAAATATGCGT
CTTCTGGATGTGAAATAACTCTGCCACACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCC
TTATTCCTGTCCGTGCCCTGGTGCTTCCTGTAAATGGCAAGGCTCTCTGGATGCTGTAATGCCCCATCTG
ATGCATCAGCATAAGTCCATTACAACCCTACAGGGAGAGGATATAGTTTTTCTTGCTACAGACATTAATC
TTCCTGGTGCTGTTGACTGGGTGATGATGCAGTCCTGTTTTGGCTTTCACTTCATGTTAGTCTTAGAGAA
ACAGGAAAAATACGATGGTCACCAGCAGTTCTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCT
GAAAATTTTGCTTACCGACTTGAGCTAAATGGTCATAGGCGACGATTGACTTGGGAAGCGACTCCTCGAT
CTATTCATGAAGGAATTGCAACAGCCATTATGAATAGCGACTGTCTAGTCTTTGACACCAGCATTGCACA
GCTTTTTGCAGAAAATGGCAATTTAGGCATCAATGTAACTATTTCCATGTGTTGAAATGGCAATCAAACA
TTTTCTGGCCAGTGTTTAAAACTTCAGTTTCACAGAAAATAAGGCACCCATCTGTCTGCCAACCTAAAAC
TCTTTCGGTAGGTGGAAGCTAGACACATGAAGGTAAATAAAAAGAAAGGCTGTTAAATACAGGAAACAGT
TGCATGTAGTAACACTAATATATTTAAAAATAAGTCAACAGTAAACCACTGAAAAAATATATGTATATAC
ACCCAAGATGGGCATCTTTTGTATTAAGAAAGGAAGCATTGTAAAATAATTCTGAGTTTTGTGTTTGTTG
TAGATTGATTGTATTGTTGAAAAAGTTTGTTTTTGCCGTGGGAGTGTGTGCCTGCGTGGGTGTGTGCGTG
TTGGTTTTTCCTTAACTGACAAGCCATCTTGAGTGGTCATGGGCCACTGCTTTTCCCTTTGTGAGTCAAT

ACATAGTGCTGCTGTGTGCTTTTTTTGTGTGTATTTGCTAATTTTTATTAATTTTAGTTTTTCATTAAAT
AAATTTGACTTTTCTGTAATTCAGGTTTTTCCTTTTTTTGTACCATTTTAAAGTTAGTATCTTTTGATAT
GCATATTTGTTTATGGTAAAAAATTTATAACGTGTTCAATATTTTCTTTTCCCCCATTAATCAGTTCATT
AGAAATATTTTAAAATCAGCTATTTTGTGAAGCCATGAGTTCCAGAAAGTAAAGGTGACATCGGAAAAAT
AATCCCAAGCTATTTAAAGCATCTATAAGGTGCTCTCTTTCTGTCTTCTACAGATGAGTCACACCTTTGA
GCTTAATCTTGAAGGTTAGAGATAAATTGATTTTTATAAATACTGCAAATCAGGCTTTTGTTTCCTTTTT
CAGATATCTTGGACAAATCACATATTTTAAAATTTGTTCTTGTATTTATTGGTTTTGCAGAAGAAGGCAT
CGTCATGCACAGTATTTGTAATTAAAAGCAAATCATTTGTTTAAAAAGGCAGTTTGCAAAAAAATGTTTTT
GGTCTTTTATAATTCTCA

Human SIAH1 mRNA sequence - var3 (public gi: 16551141) (SEQ ID NO: 137)
TTTATAATAGCCCTCCAAATGGGTGTGACGTATTTTGATTCTATGTCCCTAATGACTAGTGATGTTGAGC
ATTTTCTAGCATTGATTTTTAAGATGTTACCCAAAGACCCCTTGTATCAAAATAAGCTGGATTTTTTTAT
TGAAAATTATTAACTCTAGAAATTTTAGTTTAAACTAGACTTAGGATATGTGTATTTTACTGGTATTCC
ACGTTTTATGCATGGGTTTTTAAAACTTCTCAAGTATTAAAACTAAAAGCTTTAGGTGCTTTGCTTATCA
AGAAATCCTACACTGTCCACTGGAGACATCCATGTTTTTACTTGGCTCTGCCCCTTTAGTGGTCCCTGTG
AACCTTACCTCAAACCATGCATCTGGGGCAGAGATCCTTACTTGCTTGGTGGTTACAAATGCAAATACAG
TGAAGAATGTCATCTTTGTGATTGTTCCTGAAATAGTTCACGAGAAATCCATGACCGTAAAGTACTGTGA
TAGTGATGTCTACCACTGTGAGCTTCCAGTACTAGGTGATTGGTCTGCATTCACAGTGACCAAAATCAGC
TATGTGGCCAGGTAATTCACTGCTGAGGGCTTTGGATTTTCCTTTATGAACTACTGAAATGAGGTCAACT
TGACTATTACTAAGGGACATTTTGCTACAAAGAATGTTAGTTTTGCCAATTCCCTTTCCAAATCTAAAAT
TTATTTTAACCAGGATTTTAGATGTAAACATCAAGTAGTTTTGGTTGTTTCAATGAAGTAACATGTTTAA
GCTCACATTATTTGAAGTACTTCAGTTCCTATTGCCATGAAAATTGTATCCAGCAGCTAAAAAAAAAAAA
AAAAAAAAGACTACAGTTAGTCATTATCCAATTTGATGATTTATGGTCCAACACTAATGCTCATTTTTTT
TGTTTTTTTACAAACATTTGGTGGATACCACAATGAAAACTGCACTTAAAAAACAAAAATGCTGAAAGA
GGAAGGAAATATCAAAAAGGTCTGAATAGACAACAGGCAAATATGCTTCCACCTACCGAAAGAGTTTTAG
GTTGGCAGACAGATGGGTGCCTTATTTTCTGTGAAACTGAAGTTTTAAACACTGGCCAGAAAATGTTTGA
TTGCCATTTCAACACATGGAAATAGTTACATTGATGCCTAAATTGCCATTTTCTGCAAAAAGCTGTGCAA
TGCTGGTGTCAAAGACTAGACAGTCGCTATTCATAATGGCTGTTGCAATTCCTTCATGAATAGATCGAGG
AGTCGCTTCCCAAGTCAATCGTCGCCTATGACCATTTAGCTCAAGTCGGTAAGCAAAATTTTCAGCTTGC
TTGCGTGTTCCTATCAGCTGTACGATTGCGAAGAACTGCTGGTGACCATCGTATTTTTCCTGTTTCTCTA
AGACTAACATGAAGTGAAAGCCAAAACAGGACTGCATCATCACCCAGTCAACAGCACCAGGAAGATTAAT
GTCTGTAGCAAGAAAAACTATATCCTCTCCCTGTAGGGTTGTAATGGACTTATGCTGATGCATCAGATGG
GGCATTACAGCATCCAGAGAGCCTTGCCATTTACAGGAAGCACCAGGGCACGGACAGGAATAAGGCCTAA
ACTCACAGAGCTCTTCATGGTCTGCTTTTTCTGTGTGTGGCAGAGTTATTTCACATCCAGAAGACGCATA
TTTACAGGGGAAAAGTACTGAATTAGCCACTTTCTCCATAGCCAAGTTGCGAATGGATCCCAAAGGGCCC
CGGCAAGTTGGACAACATGTGAGCTTTGGGCGACAGTTGCTACAAACAAGATGGCCACTCTGACATTGAA
GAATGGGCGGTAACACATAGTCAAAGCAGACTGGACACTCAAAAAGACTCGCCAAGTCATTGTTGGATGC
AGTTGTGCCAGTCAGGGCAGGCACCCTCTGGGATGGTGGACACTTCGAGGTACCGGTAGGTAATGCTGTA
GCAGTCTGACGGCTCATTTCTGAAATAAATACATAAGGAGGCAGGAGAAAAATAATTATAACCATGACTT
ACTTTATAAATAATGTTTACATGCCATAAGTCCTTTTAAAGTTTCATACAAAATTTACTGAGCAAAAGAG
GAAGAAAAATAGGATTAAAAAAGATATT

Human SIAH1 mRNA sequence - var4 (public gi: 21753769) (SEQ ID NO: 138)
TTTCACCCCCAAGACAAATAGTGGCCTGCCATTTTCCAGCCCAGGTAGCTTCCTGGGAAAAGTTGCTTGT
TTTATCTTTGACTCAGCCTGGCTAGTTACATTGTCGATTATTTCTTCCAGATGATATTTACCTGTTAAAT
AATGTTTATTACTCTGCTGATGAATGTTTTCAGCAACGCTGGAGAACCCTAGGCTGCAAGGGGTTCTTCA
CCTGTTGACTCCATCCCCCACCCCCAGTATGGCATATATCTCTGCCGTGCTATCATCTTTATTCTTTCCT
TTTTCATTGTCTCTTCCTGACTGTCCTCCTTTGTTCATTATGTCTGACACATATTGTGGATTGAAAGTAG
AACAGAAAGATATACTTTCTCTACCAGACTAAAAAGTTTTGAGATGGCCCTCCATTTCTCCCATGCCTCA
CTTCACCTTAGTTGTGTTTTTATTTATTTTTATATTTTCGCCACCTTCACTAGCGAGTACATCCCCTCAC
TCTTGAGGTGGGCACTGATCAGTAGGAAATAAGATTAATACCTGGCTGGTGATAATTTGGGGGGAAGACT
TAATTAGATAGAGATGGATAATGGGATGGCAGCAGACCTTTCCCCTTGTGACCCTTCCCCTCATTTCCAA
AATACACCTCTAGAGTAGATAATTGCTTACCATTAAGAAGAGTTAATGGAAGGTGATACTCTGATTCTTT
GGCATTGGAACTACATTCAATCCGCGGTATCCTCGGATTAGTTCTAGGACCCCCTTCTCCATACCAAAAC
CTGAGGATGCTCAAGTCCCTGATAGAAAATGGTGTCATATTTGTATGTGCATATTCTCTTGTATAATTTA
AGTGATCTCTGGATTACTTAATACAATGTAAACAATATGTAAATAGTTGTTATAGACTGTATTTTAAAAA
TTTTGTATTCTTTATAAATTTTTCTGAATATTTTCAATCCATGGCTGGTGAAGTCCTCGGATGCAGACCG
TGTGGATACAGAGTGCCGATTTTATACAGGAGTTCACCTGTAACTCCCTGTACCTATCAACAGCTGACTC
CAAATTAGAAAGAAATAGAGTAAGGGAGCCTCAGGGAGAGTCCTAGCAAAACGGATTCGATTAAACTTCA
GTTCCTTGTATAGTTTCTTTAGTTGTTTATGGTCCATTTTCTATTTTAGCATTTATTATTCTATGTAGTC
TATCCAAAGACGATTAAGGGAGTTCCACATGTTTTCCGGAACATTTTGAAAAGAGAGCTTATCCAGTGTA

CAGATCCTAATAAAGTGCACATTCAGTGTAATTTTATTTTTTAATATCTTTTTAATCCTATTTTTCTT
CCTCTTTTGCTCAGTAAATTTTGTATGAAACTTTAAAAGGACTTATGGCATGTAAACATTATTTATAAAG
TAAGTCATGGTTATAATTATTTTTCTCCTGCCTCCTTATGTATTTATTTCAGAAATGAGCCGTCAGACTG
CTACAGCATTACCTACCGGTACCTCGAAGTGTCCACCATCCCAGAGGGTGCCTGCCCTGACTGGCACAAC
TGCATCCAACAATGACTTGGCGAGTCTTTTTGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCCATT
CTTCAATGTCAGAGTGGCCATCTTGTTTGTAGCAACTGTCGCCCAAAGCTCACATGTTGTCCAACTTGCC
GGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATTCAGTACTTTTCCCCTGTAA
ATATGCGTCTTCTGGATGTGAAATAACTCTGCCACACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAG
TTTAGGCCTTATTCCTGTCCGTGCCCTGGTGCTTCCTGTAAATGGCAAGGCTCTCTGGATGCTGTAATGC
CCCATCTGATGCATCAGCATAAGTCCATTACAACCCTACAGGGAGAGGATATAGTTTTTCTTGCTACAGA
CATTAATCTTCCTGGTGCTGTTGACTGGGTGATGATGCAGTCCTGTTTTGGCTTTCACTTCATGTTAGTC
TTAGAGAAACAGGAAAAATACGATGGTCACCAGCAGTTCTTCGCAATCGTACAGCTGATAGGAACACGCA
AGCAAGCTGAAAATTTTGCTTACCGACTTGAGCTAAATGGTCATAGGCGACGATTGACTTGGGAAGCGAC
TCCTCGATCTATTCATGAAGGAATTGCAACAGCCATTATGAATAGCGACTGTCTAGTCTTTGCCACCAGC
ATTGCACAGCTTTTTGCAGAAAATGGCAATTTAGGCATCAATGTAACTATTTCCATGTGTTGAAATGGCA
ATCAAACATTTTCTGGCCAGTGTTTAAAACTTCAGTTTCACAGAAAATAAGGCACCCATCTGTCTGCCAA
CCTAAAACTCTTTCGGTAGGTGGAAGCTAGACACATGAAGGTAAATAAAAAGAAAGGCTGTTAAATACAG
GAAACAGTTGCATGTAGTAACACTAATATATTTAAAAATAAGTCAACAGTAAACCACTGAAAAAATATAT
GTATATACACCCAAGATGGGCATCTTTTGTATTAAGAAAGGAAGCATTGTAAAATAATTCTGAGTTTTGT
GTTTGTTGTAGATTGATTGTATTGTTGAAAAAGTTTGTTTTTGCGTGGGAGTGTGTGCCTGCGTGGGTGT
GTGCGTGTTTGGGTTTTTTTCCTTTAACTGACAAGCCATCTTGAGTGGTCATGGGCCACTGCTTTTCCCT
TTGTGAGTCAATACATAGTGCTGCTGTGTGCTTTTTTTGTGTGTATTTGCTAATTTTTATTAATTTTAGT
TTTTCATTAAATAAATTTGACTTTTCTGT

Human SIAH1 mRNA sequence - var5 (public gi: 3041824) (SEQ ID NO: 139)
ATGAGCCGTCAGACTGCTACAGCATTACCTACCGGTACCTCGAAGTGTCCACCATCCCAGAGGGTGCCTG
CCCTGACTGGCACAACTGCATCCAACAATGACTTGGCGAGTCTTTTTGAGTGTCCAGTCTGCTTTGACTA
TGTGTTACCGCCCATTCTTCAATGTCAGAGTGGCCATCTTGTTTGTAGCAACTGTCGCCCAAAGCTCACA
TGTTGTCCAACTTGCCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATTCAG
TACTTTTCCCCTGTAAATATGCGTCTTCTGGATGTGAAATAACTCTGCCACACACAGAAAAAGCAGACCA
TGAAGAGCTCTGTGAGTTTAGGCCTTATTCCTGTCCGTGCCCTGGTGCTTCCTGTAAATGGCAAGGCTCT
CTGGATGCTGTAATGCCCCATCTGATGCATCAGCATAAGTCCATTACAACCCTACAGGGAGAGGATATAG
TTTTTCTTGCTACAGACATTAATCTTCCTGGTGCTGTTGACTGGGTGATGATGCAGTCCTGTTTTGGCTT
TCACTTCATGTTAGTCTTAGAGAAACAGGAAAAATACGATGGTCACCAGCAGTTCTTCGCAATCGTACAG
CTGATAGGAACACGCAAGCAAGCTGAAAATTTTGCTTACCGACTTGAGCTAAATGGTCATAGGCGACGAT
TGACTTGGGAAGCGACTCCTCGATCTATTCATGAAGGAATTGCAACAGCCATTATGAATAGCGACTGTCT
AGTCTTTGACACCAGCATTGCACAGCTTTTTGCAGAAAATGGCAATTTAGGCATCAATGTAACTATTTCC
ATGTGTTGAAATGGCAATCAAACATTTTCTGGCCAGTGTTTAAAACTTCAGTTTCACAGAAAATAAGGCA
CCCATCTGTCTGCCAACCTAAAACTCTTTCGGTAGGTAGAAGCTAGACACATGAAGGTAAATAAAAAGAA
AGGCTGTTAAATACAGGAAACAGTTCCATGTAGTAACACTAATATATTTAAAAATAAGTCAACAGTAAAC
CACTGAAAAAATATATGTATATACACCCAAGATGGGCATCTTTTGTATTAAGAAAGGAAGCATTGTAAAA
TAATTCTGAGTTTTGTGTTTGTTGTAGATTGATTGTATTGTTGAAAAAGTTTGTTTTTGCGTGGGAGTGT
GTGCCTGCGTGGGTGTGTGCGTGTTTGGGTTTTTTTCCTTTAACTGACAAGCCATCTTGAGTGGTCATGG
GCCACTGCTTTTCCCTTTGTGAGTCAATACATAGTGCTGCTGTAAGCCGTTTTTGTGTGTATTTGCTAAT
TTTTATTAATTTTAGTTTTTCATTAAATAAATTTGACTTTTCTGTAATTCAGGTTTTTCCTTTTTTTTGTA
CCATTTTAAAGTTAGTATCTTTTGATATGGCATATTTGTTTATGGTAAAAAATTTATAACGGGTTCAATA
TTTTCTTTTCCCCCATTAATCAAGTCCATTGGAAATATTTTAAAACCAGCCTATTTTGGTGAACCCATGA
GTTCCCAGAAAGTAAAGGTGACACCCGGAAAAATAATCCAAAAGCCTATTTAAAGCCACCTATAAGGTGC
CCCCCTTTCCTGTCTTCCTACAGATGAGTCACACCTTTGAGCCTTAACCTTTGAAAGGTTAGAGAATAAA
TTGATTTTTATAAATACTGCAAATCCAGGCTTTTGTTTCCTTTTTCCAGATATCCTTGGACAAATCACAT
ATTTTAAAATTTGTTCTTGTATTTATTGGTTTTGCAGAAGAAGGCATCGTCATGCACAGTATTTGTAATT
AAAAGCAAATTCATTTGTTTAAAAAGGCAGTTTGCAAAAAATGTTTTTGGTCTTTTATAATTCTCA

Human SIAH1 mRNA sequence - var6 (public gi: 17390431) (SEQ ID NO: 140)
CGGCGCCGGGAAGAGGCGGTGGCGCTGCCCGCGGTGGCGGGGGTTGGCGACGGAGCGCGTTGGTGCCAGG
ACCGGGGTCCGAGGCGCGCTCTCCGCCCACAGAAATGAGCCGTCAGACTGCTACAGCATTACCTACCGGT
ACCTCGAAGTGTCCACCATCCCAGAGGGTGCCTGCCCTGACTGGCACAACTGCATCCAACAATGACTTGG
CGAGTCTTTTTGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCCATTCTTCAATGTCAGAGTGGCCA
TCTTGTTTGTAGCAACTGTCGCCCAAAGCTCACATGTTGTCCAACTTGCCGGGGCCCTTTGGGATCCATT
CGCAACTTGGCTATGGAGAAAGTGGCTAATTCAGTACTTTTCCCCTGTAAATATGCGTCTTCTGGATGTG
AAATAACTCTGCCACACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCCTTATTCCTGTCC
GTGCCCTGGTGCTTCCTGTAAATGGCAAGGCTCTCTGGATGCTGTAATGCCCCATCTGATGCATCAGCAT

AAGTCCATTACAACCCTACAGGGAGAGGATATAGTTTTTCTTGCTACAGACATTAATCTTCCTGGTGCTG
TTGACTGGGTGATGATGCAGTCCTGTTTTGGCTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAAAATA
CGATGGTCACCAGCAGTTCTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGAAAATTTTGCT
TACCGACTTGAGCTAAATGGTCATAGGCGACGATTGACTTGGGAAGCGACTCCTCGATCTATTCATGAAG
GAATTGCAACAGCCATTATGAATAGCGACTGTCTAGTCTTTGACACCAGCATTGCACAGCTTTTTTGCAG
AAAATGGCAATTTAGGCATCAATGTAACTATTTCCATGTGTTGAAATGGCAATCAAACATTTTCTGGCCA
GTGTTTAAAACTTCAGTTTCACAGAAAATAAGGCACCCATCTGTCTGCCAACCTAAAACTCTTTCGGTAG
GTGGAAGCTAGACACATGAAGGTAAATAAAAAGAAAGGCTGTTAAATACAGGAAACAGTTGCATGTAGTA
ACACTAATATATTTAAAAATAAGTCAACAGTAAACCACTGAAAAAATATATGTATATACACCCAAGATGG
GCATCTTTTGTATTAAGAAAGGAAGCATTGTAAAATAATTCTGAGTTTTGTGTTTGTTGTAGATTGATTG
TATTGTTGAAAAAGTTTGTTTTTGCGTGGGAGTGTGTGCCTGCGTGGGTGTGTGCGTGTTTGGGTTTTTT
TCCTTTAACTGACAAGCCATCTTGAGTGGTCATGGGCCACTGCTTTTCCCTTTGTGAGTCAATACATAGT
GCTGCTGTGTGCTTTTTTTGTGTGTATTTGCTAATTTTTATTAATTTTAGTTTTTCATTAAATAAATTTG
ACTTTTCTGTAATTCAGGTTTTTCCTTTTTTTGTACCATTTTAAAGTTAGTATCTTTTGATATGCATATT
TGTTTATGGTAAAAAATTTATAACGTGTTCAATATTTTCTTTTCCCCCATTAATCAGTTCATTAGAAATA
TTTTAAAATCAGCTATTTTGTGAAGCCATGAGTTCCAGAAAGTAAAGGTGACATCGGAAAAATAATCAAA
AGCTATTTAAAGCATCTATAAGGTGCTCTCTTTCTGTCTTCTACAGATGAGTCACACCTTTGAGCTTAAT
CTTTGAAAGGTTAGAGAATAAATTGATTTTTATAAATACTGCAAATCAGGCTTTTGTTTCCTTTTTCAGA
TATCTTGGACAAATCACATATTTTAAAATTTGTTCTTGTATTTATTGGTTTTGCAGAAGAAGGCATCGTC
ATGCACAGTATTTGTAATTAAAAGCAAATCATTTGTTTAAAAAGGCAGTTTGCAAAAAATGTTTTTGGTC
TTTTATAATTCTCATTAAAAGAATATCTGGCCATTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAA

Human SIAH1 mRNA sequence - var7 (public gi: 23274141) (SEQ ID NO: 141)
GTCCCGTCGGTCTCGGCGCCGGGAAGAGGCGGTGGCGCTGCCCGCGGTGGCGGGGGTTGGCGACGGAGCG
CGTTGGTGCCAGGACCGGGGTCCGAGGCGCGCTCTCCGCCCACAGAAATGAGCCGTCAGACTGCTACAGC
ATTACCTACCGGTACCTCGAAGTGTCCACCATCCCAGAGGGTGCCTGCCCTGACTGGCACAACTGCATCC
AACAATGACTTGGCGAGTCTTTTTGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCCATTCTTCAAT
GTCAGAGTGGCCATCTTGTTTGTAGCAACTGTCGCCCAAAGCTCACATGTTGTCCAACTTGCCGGGGCCC
TTTGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATTCAGTACTTTTCCCCTGTAAATATGCG
TCTTCTGGATGTGAAATAACTCTGCCACACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGC
CTTATTCCTGTCCGTGCCCTGGTGCTTCCTGTAAATGGCAAGGCTCTCTGGATGCTGTAATGCCCCATCT
GATGCATCAGCATAAGTCCATTACAACCCTACAGGGAGAGGATATAGTTTTTCTTGCTACAGACATTAAT
CTTCCTGGTGCTGTTGACTGGGTGATGATGCAGTCCTGTTTTGGCTTTCACTTCATGTTAGTCTTAGAGA
AACAGGAAAAATACGATGGTCACCAGCAGTTCTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGC
TGAAAATTTTGCTTACCGACTTGAGCTAAATGGTCATAGGCGACGATTGACTTGGGAAGCGACTCCTCGA
TCTATTCATGAAGGAATTGCAACAGCCATTATGAATAGCGACTGTCTAGTCTTTGACACCAGCATTGCAC
AGCTTTTTGCAGAAAATGGCAATTTAGGCATCAATGTAACTATTTCCATGTGTTGAAATGGCAATCAAAC
ATTTCTGGCCAGTGTTTAAAACTTCAGTTTCACAGAAAATAAGGCACCCATCTGTCTGCCAACCTAAAA
CTCTTTCGGTAGGTGGAAGCTAGACACATGAAGGTAAATAAAAAGAAAGGCTGTTAAATACAGGAAACAG
TTGCATGTAGTAACACTAATATATTTAAAAATAAGTCAACAGTAAACCACTGAAAAAATATATGTATATA
CACCCAAGATGGGCATCTTTTGTATTAAGAAAGGAAGCATTGTAAAATAATTCTGAGTTTTGTGTTTGTT
GTAGATTGATTGTATTGTTGAAAAAGTTTGTTTTTGCGTGGGAGTGTGTGCCTGCGTGGGTGTGTGCGTG
TTTGGGTTTTTTTCCTTTAACTGACAAGCCATCTTGAGTGGTCATGGGCCACTGCTTTTCCCTTTGTGAG
TCAATACATAGTGCTGCTGTGTGCTTTTTTTGTGTGTATTTGCTAATTTTTATTAATTTTAGTTTTTCAT
TAAATAAATTTGACTTTTCTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human SIAH1 Protein sequence - var1 (public gi: 27503514) (SEQ ID NO: 271)
MTGKATPPSLYSWRGVLFTCLPAARTRKRKEMSRQTATALPTGTSKCPPSQRVPALTGTTASNNDLASLF
ECPVCFDYVLPPILQCQSGHLVCSNCRPKLTCCPTCRGPLGSIRNLAMEKVANSVLFPCKYASSGCEITL
PHTEKADHEELCEFRPYSCPCPGASCKWQGSLDAVMPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWV
MMQSCFGFHFMLVLEKQEKYDGHQQFFAIVQLIGTRKQAENFAYRLELNGHRRRLTWEATPRSIHEGIAT
AIMNSDCLVFDTSIAQLFAENGNLGINVTISMC

Human SIAH1 Protein sequence - var2 (public gi: 4506947) (SEQ ID NO: 272)
MSRQTATALPTGTSKCPPSQRVPALTGTTASNNDLASLFECPVCFDYVLPPILQCQSGHLVCSNCRPKLT
CCPTCRGPLGSIRNLAMEKVANSVLFPCKYASSGCEITLPHTEKADHEELCEFRPYSCPCPGASCKWQGS
LDAVMPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWVMMQSCFGFHFMLVLEKQEKYDGHQQFFAIVQ
LIGTRKQAENFAYRLELNGHRRRLTWEATPRSIHEGIATAIMNSDCLVFDTSIAQLFAENGNLGINVTIS
MC

Unigene Name: SMN1 **Unigene ID**: Hs.288986 **Clone ID**: GD_1114

Human SMN1 mRNA sequence - var1 (public gi: 624185) (SEQ ID NO: 142)
CGGGGCCCCACGCTGCGCATCCGCGGGTTTGCTATGGCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCC
CGGAGCAGGAGGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTCTGACATTTGGGATGA
TACAGCACTGATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTAAAGAATGGTGACATT
TGTGAAACTTCGGGTAAACCAAAAACCACACCTAAAAGAAAACCTGCTAAGAAGAATAAAAGCCAAAAGA
AGAATACTGCAGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCATTTGGTCAGAAGACGG
TTGCATTTACCCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTTGTGGTTTACACTGGA
TATGGAAATAGAGAGGAGCAAAATCTGTCCGATCTACTTTCCCCAATCTGTGAAGTAGCTAATAATATAG
AACAGAATGCTCAAGAGAATGAAAATGAAAGCCAAGTTTCAACAGATGAAAGTGAGAACTCCAGGTCTCC
TGGAAATAAATCAGATAACATCAAGCCCAAATCTGCTCCATGGAACTCTTTTCTCCCTCCACCACCCCCC
ATGCCAGGGCCAAGACTGGGACCAGGAAAGCCAGGTCTAAAATTCAATGGCCCACCACCGCCACCGCCAC
CACCACCACCCCACTTACTATCATGCTGGCTGCCTCCATTTCCTTCTGGACCACCAATAATTCCCCCACC
ACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGTATGTTAATTTCATGGTACATG
AGTGGCTATCATACTGGCTATTATATGGGTTTCAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCT
TAAATTAAGGAGAAATGCTGGCATAGAGCAGCACTAAATGACACCACTAAAGAAACGATCAGACAGATCT
GGAATGTGAAGCGTTATAGAAGATAACTGGCCTCATTTCTTCAAAATATCAAGTGTTGGGAAAGAAAAAA
GGAAGTGGAATGGGTAACTCTTCTTGATTAAAAGTTATGTAATAACCAAATGCAATGTGAAATATTTTAC
TGGACTCTTTTGAAAAACCATCTGTAAAAGACTGGGGTGGGGGTGGGAGGCCAGCACGGTGGTGAGGCAG
TTGAGAAAATTTGAATGTGGATTAGATTTTGAATGATATTGGATAATTATTGGTAATTTTATGGCCTGTG
AGAAGGGTGTTGTAGTTTATAAAAGACTGTCTTAATTTGCATACTTAAGCATTTAGGAATGAAGTGTTAG
AGTGTCTTAAAATGTTTCAAATGGTTTAACAAAATGTATGTGAGGCGTATGTGGCAAAATGTTACAGAAT
CTAACTGGTGGACATGGCTGTTCATTGTACTGTTTTTTTCTATCTTCTATATGTTTAAAAGTATATAATA
AAAATATTTAATTTTTTTTTA

Human SMN1 mRNA sequence - var2 (public gi: 15929773) (SEQ ID NO: 143)
GGCCCCACGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCCCGG
AGCAGGAGGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTCTGACATTTGGGATGATAC
AGCACTGATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTAAAGAATGGTGACATTTGT
GAAACTTCGGGTAAACCAAAAACCACACCTAAAAGAAAACCTGCTAAGAAGAATAAAAGCCAAAAGAAGA
ATACTGCAGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCATTTGGTCAGAAGACGGTTG
CATTTACCCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTTGTGGTTTACACTGGATAT
GGAAATAGAGAGGAGCAAAATCTGTCCGATCTACTTTCCCCAATCTGTGAAGTAGCTAATAATATAGAAC
AGAATGCTCAAGAGAATGAAAATGAAAGCCAAGTTTCAACAGATGAAAGTGAGAACTCCAGGTCTCCTGG
AAATAAATCAGATAACATCAAGCCCAAATCTGCTCCATGGAACTCTTTTCTCCCTCCACCACCCCCCATG
CCAGGGCCAAGACTGGGACCAGGAAAGCCAGGTCTAAAATTCAATGGCCCACCACCGCCACCGCCACCAC
CACCACCCCACTTACTATCATGCTGGCTGCCTCCATTTCCTTCTGGACCACCAATAATTCCCCCACCACC
TCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGTATGTTAATTTCATGGTACATGAGT
GGCTATCATACTGGCTATTATATGGGTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAA
ATTAAGGAGAAATGCTGGCATAGAGCAGCACTAAATGACACCACTAAAGAAACGATCAGACAGATCTGGA
ATGTGAAGCGTTATAGAAGATAACTGGCCTCATTTCTTCAAAATATCAAGTGTTGGGAAAGAAAAAAGGA
AGTGGAATGGGTAACTCTTCTTGATTAAAAGTTATGTAATAACCAAATGCAATGTGAAATATTTTACTGG
ACTCTATTTTGAAAAACCATCTGTAAAAGACTGAGGTGGGGGTGGGAGGCCAGCACGGTGGTGAGGCAGT
TGAGAAAATTTGAATGTGGATTAGATTTTGAATGATATTGGATAATTATTGGTAATTTTATGAGCTGTGA
GAAGGGTGTTGTAGTTTATAAAAGACTGTCTTAATTTGCATACTTAAGCATTTAGGAATGAAGTGTTAGA
GTGTCTTAAAATGTTTCAAATGGTTTAACAAAATGTATGTGAGGCGTATGTGGCAAAATGTTACAGAATC
TAACTGGTGGACATGGCTGTTCATTGTACTGTTTTTTTCTATCTTCTATATGTTTAAAAGTATATAATAA
AAATATTTAATTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human SMN1 mRNA sequence - var3 (public gi: 13259511) (SEQ ID NO: 144)
CCACAAATGTGGGAGGGCGATAACCACTCGTAGAAAGCGTGAGAAGTTACTACAAGCGGTCCTCCCGGCC
ACCGTACTGTTCCGCTCCCAGAAGCCCCGGGCGGCGGAAGTCGTCACTCTTAAGAAGGGACGGGGCCCCA
CGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCCCGGAGCAGGA
GGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTCTGACATTTGGGATGATACAGCACTG
ATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTAAAGAATGGTGACATTTGTGAAACTT
CGGGTAAACCAAAAACCACACCTAAAAGAAAACCTGCTAAGAAGAATAAAAGCCAAAAGAAGAATACTGC
AGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCATTTGGTCAGAAGACGGTTGCATTTAC
CCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTTGTGGTTTACACTGGATATGGAAATA
GAGAGGAGCAAAATCTGTCCGATCTACTTTCCCCAATCTGTGAAGTAGCTAATAATATAGAACAGAATGC

TCAAGAGAATGAAAATGAAAGCCAAGTTTCAACAGATGAAAGTGAGAACTCCAGGTCTCCTGGAAATAAA
TCAGATAACATCAAGCCCAAATCTGCTCCATGGAACTCTTTTCTCCCTCCACCACCCCCCATGCCAGGGC
CAAGACTGGGACCAGGAAAGATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGA
TGCTTTGGGAAGTATGTTAATTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGGGTTTCAGA
CAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATTAAGGAGAAATGCTGGCATAGAGCAGCACTA
AATGACACCACTAAAGAAACGATCAGACAGATCTGGAATGTGAAGCGTTATAGAAGATAACTGGCCTCAT
TTCTTCAAAATATCAAGTGTTGGGAAAGAAAAAAGGAAGTGGAATGGGTAACTCTTCTTGATTAAAAGTT
ATGTAATAACCAAATGCAATGTGAAATATTTTACTGGACTCTTTTGAAAAACCATCTGTAAAAGACTGGG
GTGGGGGTGGGAGGCCAGCACGGTGGTGAGGCAGTTGAGAAAATTTGAATGTGGATTAGATTTTGAATGA
TATTGGATAATTATTGGTAATTTTATGGCCTGTGAGAAGGGTGTTGTAGTTTATAAAAGACTGTCTTAAT
TTGCATACTTAAGCATTTAGGAATGAAGTGTTAGAGTGTCTTAAAATGTTTCAAATGGTTTAACAAAATG
TATGTGAGGCGTATGTGGCAAAATGTTACAGAATCTAACTGGTGGACATGGCTGTTCATTGTACTGTTTT
TTTCTATCTTCTATATGTTTAAAAGTATATAATAAAAATATTTAATTTTTTTTTA

Human SMN1 mRNA sequence - var4 (public gi: 13111817) (SEQ ID NO: 145)
GGGGCCCCACGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCCC
GGAGCAGGAGGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTCTGACATTTGGGATGAT
ACAGCACTGATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTAAAGAATGGTGACATTT
GTGAAACTTCGGGTAAACCAAAAACCACACCTAAAAGAAAACCTGCTAAGAAGAATAAAAGCCAAAAGAA
GAATACTGCAGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCATTTGGTCAGAAGACGGT
TGCATTTACCCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTTGTGGTTTACACTGGAT
ATGGAAATAGAGAGGAGCAAAATCTGTCCGATCTACTTTCCCCAATCTGTGAAGTAGCTAATAATATAGA
ACAGAATGCTCAAGAGAATGAAAATGAAAGCCAAGTTTCAACAGATGAAAGTGAGAACTCCAGGTCTCCT
GGAAATAAATCAGATAACATCAAGCCCAAATCTGCTCCATGGAACTCTTTTCTCCCTCCACCACCCCCCA
TGCCAGGGCCAAGACTGGGACCAGGAAAGCCAGGTCTAAAATTCAATGGCCCACCACCGCCACCGCCACC
ACCACCACCCCACTTACTATCATGCTGGCTGCCTCCATTTCCTTCTGGACCACCAATAATTCCCCCACCA
CCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGTATGTTAATTTCATGGTACATGA
GTGGCTATCATACTGGCTATTATATGGAAATGCTGGCATAGAGCAGCACTAAATGACACCACTAAAGAAA
CGATCAGACAGATCTGGAATGTGAAGCGTTATAGAAGATAACTGGCCTCATTTCTTCAAAATATCAAGTG
TTGGGAAAGAAAAAAGGAAGTGGAATGGGTAACTCTTCTTGATTAAAAGTTATGTAATAACCAAATGCAA
TGTGAAATATTTTACTGGACTCTATTTTGAAAAACCATCTGTAAAAGACTGAGGTGGGGGTGGGAGGCCA
GCACGGTGGTGAGGCAGTTGAGAAAATTTGAATGTGGATTAGATTTTGAATGATATTGGATAATTATTGG
TAATTTTATGAGCTGTGAGAAGGGTGTTGTAGTTTATAAAAGACTGTCTTAATTTGCATACTTAAGCATT
TAGGAATGAAGTGTTAGAGTGTCTTAAAATGTTTCAAATGGTTTAACAAAATGTATGTGAGGCGTATGTG
GCAAAATGTTACAGAATCTAACTGGTGGACATGGCTGTTCATTGTACTGTTTTTTTCTATCTTCTATATG
TTTAAAAGTATATAATAAAAATATTTAATTTTTTTTTAAAAA

Human SMN1 mRNA sequence - var5 (public gi: 13259515) (SEQ ID NO: 146)
CCACAAATGTGGGAGGGCGATAACCACTCGTAGAAAGCGTGAGAAGTTACTACAAGCGGTCCTCCCGGCC
ACCGTACTGTTCCGCTCCCAGAAGCCCCGGGCGGCGGAAGTCGTCACTCTTAAGAAGGGACGGGGCCCCA
CGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCCCGGAGCAGGA
GGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTCTGACATTTGGGATGATACAGCACTG
ATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTAAAGAATGGTGACATTTGTGAAACTT
CGGGTAAACCAAAAACCACACCTAAAAGAAAACCTGCTAAGAAGAATAAAAGCCAAAAGAAGAATACTGC
AGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCATTTGGTCAGAAGACGGTTGCATTTAC
CCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTTGTGGTTTACACTGGATATGGAAATA
GAGAGGAGCAAAATCTGTCCGATCTACTTTCCCCAATCTGTGAAGTAGCTAATAATATAGAACAGAATGC
TCAAGAGAATGAAAATGAAAGCCAAGTTTCAACAGATGAAAGTGAGAACTCCAGGTCTCCTGGAAATAAA
TCAGATAACATCAAGCCCAAATCTGCTCCATGGAACTCTTTTCTCCCTCCACCACCCCCCATGCCAGGGC
CAAGACTGGGACCAGGAAAGCCAGGTCTAAAATTCAATGGCCCACCACCGCCACCGCCACCACCACCACC
CCACTTACTATCATGCTGGCTGCCTCCATTTCCTTCTGGACCACCAATAATTCCCCCACCACCTCCCATA
TGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGTATGTTAATTTCATGGTACATGAGTGGCTATC
ATACTGGCTATTATATGGGTTTCAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATTAAGG
AGAAATGCTGGCATAGAGCAGCACTAAATGACACCACTAAAGAAACGATCAGACAGATCTGGAATGTGAA
GCGTTATAGAAGATAACTGGCCTCATTTCTTCAAAATATCAAGTGTTGGGAAAGAAAAAAGGAAGTGGAA
TGGGTAACTCTTCTTGATTAAAAGTTATGTAATAACCAAATGCAATGTGAAATATTTTACTGGACTCTTT
TGAAAAACCATCTGTAAAAGACTGGGGTGGGGGTGGGAGGCCAGCACGGTGGTGAGGCAGTTGAGAAAAT
TTGAATGTGGATTAGATTTTGAATGATATTGGATAATTATTGGTAATTTTATGGCCTGTGAGAAGGGTGT
TGTAGTTTATAAAAGACTGTCTTAATTTGCATACTTAAGCATTTAGGAATGAAGTGTTAGAGTGTCTTAA
AATGTTTCAAATGGTTTAACAAAATGTATGTGAGGCGTATGTGGCAAAATGTTACAGAATCTAACTGGTG
GACATGGCTGTTCATTGTACTGTTTTTTTCTATCTTCTATATGTTTAAAAGTATATAATAAAAATATTTA
ATTTTTTTTTA

Human SMN1 Protein sequence - var1 (public gi: 13259512) (SEQ ID NO: 273)
MAMSSGGSGGGVPEQEDSVLFRRGTGQSDDSDIWDDTALIKAYDKAVASFKHALKNGDICETSGKPKTTP
KRKPAKKNKSQKKNTAASLQQWKVGDKCSAIWSEDGCIYPATIASIDFKRETCVVVYTGYGNREEQNLSD
LLSPICEVANNIEQNAQENENESQVSTDESENSRSPGNKSDNIKPKSAPWNSFLPPPPPMPGPRLGPGKI
IPPPPPICPDSLDDADALGSMLISWYMSGYHTGYYMGFRQNQKEGRCSHSLN

Human SMN1 Protein sequence - var2 (public gi: 12654181) (SEQ ID NO: 274)
MAMSSGGSGGGVPEQEDSVLFRRGTGQSDDSDIWDDTALIKAYDKAVASFKHALKNGDICETSGKPKTTP
KRKPAKKNKSQKKNTAASLQQWKVGDKCSAIWSEDGCIYPATIASIDFKRETCVVVYTGYGNREEQNLSD
LLSPICEVANNIEQNAQENENESQVSTDESENSRSPGNKSDNIKPKSAPWNSFLPPPPPMPGPRLGPGKP
GLKFNGPPPPPPPPPPHLLSCWLPPFPSGPPIIPPPPPICPDSLDDADALGSMLISWYMSGYHTGYYMEM
LA

Human SMN1 Protein sequence - var3 (public gi: 4507091) (SEQ ID NO: 275)
MAMSSGGSGGGVPEQEDSVLFRRGTGQSDDSDIWDDTALIKAYDKAVASFKHALKNGDICETSGKPKTTP
KRKPAKKNKSQKKNTAASLQQWKVGDKCSAIWSEDGCIYPATIASIDFKRETCVVVYTGYGNREEQNLSD
LLSPICEVANNIEQNAQENENESQVSTDESENSRSPGNKSDNIKPKSAPWNSFLPPPPPMPGPRLGPGKP
GLKFNGPPPPPPPPPPHLLSCWLPPFPSGPPIIPPPPPICPDSLDDADALGSMLISWYMSGYHTGYYMGF
RQNQKEGRCSHSLN

Human SMN2 mRNA sequence - var1 (public gi: 736410) (SEQ ID NO: 147)
GCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCCCGGAGCAGGAGGATTCCGTGCTGTTCCGGCGCGGCA
CAGGCCAGAGCGATGATTCTGACATTTGGGATGATACAGCACTGATAAAAGCATATGATAAAGCTGTGGC
TTCATTTAAGCATGCTCTAAAGAATGGTGACATTTGTGAAACTTCGGGTAAACCAAAAACCACACCTAAA
AGAAAACCTGCTAAGAAGAATAAAAGCCAAAAGAAGAATACTGCAGCTTCCTTACAACAGTGGAAAGTTG
GGGACAAATGTTCTGCCATTTGGTCAGAAGACGGTTGCATTTACCCAGCTACCATTGCTTCAATTGATTT
TAAGAGAGAAACCTGTGTTGTGGTTTACACTGGATATGGAAATAGAGAGGAGCAAAATCTGTCCGATCTA
CTTTCCCCAATCTGTGAAGTAGCTAATAATATAGAACAGAATGCTCAAGAGAATGAAAATGAAAGCCAAG
TTTCAACAGATGAAAGTGAGAACTCCAGGTCTCCTGGAAATAAATCAGATAACATCAAGCCCAAATCTGC
TCCATGGAACTCTTTTCTCCCTCCACCACCCCCCATGCCAGGGCCAAGACTGGGACCAGGAAAGCCAGGT
CTAAAATTCAATGGCCCACCACCGCCACCGCCACCACCACCACCCCACTTACTATCATGCTGGCTGCCTC
CATTTCCTTCTGGACCACCAATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGA
TGCTTTGGGAAGTATGTTAATTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGGGTTTTAGA
CAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATTAAGGAGAAATGCTGGCATAGAGCAGCACTA
AATGACACCACTAAAGAAACGATCAGACAGATCTGGAATGTGAAGCGTTATAGAAGATAACTGGCCTCAT
TTCTTCAAAATATCAAGTGTTGGGAAAGAAAAAAGGAAGTGGAATGGGTAACTCTTCTTGATTAAAAGTT
ATGTAATAACCAAATGCAATGTGAAATATTTTACTGGACTCTATTTTGAAAAACCATCTGTAAAAGACTG
AGGTGGGGGTGGGAGGCCAGCACGGTGGTGAGGCAGTTGAGAAAATTTGAATGTGGATTAGACTTTGAAT
GATATTGGATAATTATTGGTAATTTTATGAGCTGTGAGAAGGGTGTTGTAGTTTATAAAAGACTGTCTTA
ATTTGCATACTTAAGCATTTAGGAATGAAGTGTTAGAGTGTCTTAAAATGTTTCAAATGGTTTAACAAAA
TGTATGTGAGGCGTATGTGGCAAAATGTTACAGAATCTAACTGGTGGACATGGCTGTTCATTGTACTGTT
TTTTTCTATCTTCTATATGTTTAAAAGTATATAATAAAAATATTTAATTTTTTTTAAAAAAAAAA

Human SMN2 mRNA sequence - var2 (public gi: 13259530) (SEQ ID NO: 148)
CCACAAATGTGGGAGGGCGATAACCACTCGTAGAAAGCGTGAGAAGTTACTACAAGCGGTCCTCCCGGCC
ACCGTACTGTTCCGCTCCCAGAAGCCCCGGGCGGCGGAAGTCGTCACTCTTAAGAAGGGACGGGGCCCCA
CGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCCCGGAGCAGGA
GGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTCTGACATTTGGGATGATACAGCACTG
ATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTAAAGAATGGTGACATTTGTGAAACTT
CGGGTAAACCAAAAACCACACCTAAAAGAAAACCTGCTAAGAAGAATAAAAGCCAAAAGAAGAATACTGC
AGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCATTTGGTCAGAAGACGGTTGCATTTAC
CCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTTGTGGTTTACACTGGATATGGAAATA
GAGAGGAGCAAAATCTGTCCGATCTACTTTCCCCAATCTGTGAAGTAGCTAATAATATAGAACAGAATGC
TCAAGAGAATGAAAATGAAAGCCAAGTTTCAACAGATGAAAGTGAGAACTCCAGGTCTCCTGGAAATAAA
TCAGATAACATCAAGCCCAAATCTGCTCCATGGAACTCTTTTCTCCCTCCACCACCCCCCATGCCAGGGC
CAAGACTGGGACCAGGAAAGATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGA
TGCTTTGGGAAGTATGTTAATTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGGAAATGCTG
GCATAGAGCAGCACTAAATGACACCACTAAAGAAACGATCAGACAGATCTGGAATGTGAAGCGTTATAGA
AGATAACTGGCCTCATTTCTTCAAAATATCAAGTGTTGGGAAAGAAAAAAGGAAGTGGAATGGGTAACTC
TTCTTGATTAAAAGTTATGTAATAACCAAATGCAATGTGAAATATTTTACTGGACTCTTTTGAAAAACCA

TCTGTAAAAGACTGAGGTGGGGGTGGGAGGCCAGCACGGTGGTGAGGCAGTTGAGAAAATTTGAATGTGG
ATTAGATTTTGAATGATATTGGATAATTATTGGTAATTTTATGGCCTGTGAGAAGGGTGTTGTAGTTTAT
AAAAGACTGTCTTAATTTGCATACTTAAGCATTTAGGAATGAAGTGTTAGAGTGTCTTAAAATGTTTCAA
ATGGTTTAACAAAATGTATGTGAGGCGTATGTGGCAAAATGTTACAGAATCTAACTGGTGGACATGGCTG
TTCATTGTACTGTTTTTTTCTATCTTCTATATGTTTAAAAGTATATAATAAAAAATATTTAATTTTTTTTT
AAA

Human SMN2 mRNA sequence - var3 (public gi: 13259528) (SEQ ID NO: 149)
CCACAAATGTGGGAGGGCGATAACCACTCGTAGAAAGCGTGAGAAGTTACTACAAGCGGTCCTCCCGGCC
ACCGTACTGTTCCGCTCCCAGAAGCCCCGGGCGGCGGAAGTCGTCACTCTTAAGAAGGGACGGGGCCCCA
CGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCCCGGAGCAGGA
GGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTCTGACATTTGGGATGATACAGCACTG
ATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTAAAGAATGGTGACATTTGTGAAACTT
CGGGTAAACCAAAAACCACACCTAAAAGAAAACCTGCTAAGAAGAATAAAAGCCAAAAGAAGAATACTGC
AGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCATTTGGTCAGAAGACGGTTGCATTTAC
CCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTTGTGGTTTACACTGGATATGGAAATA
GAGAGGAGCAAAATCTGTCCGATCTACTTTCCCCAATCTGTGAAGTAGCTAATAATATAGAACAGAATGC
TCAAGAGAATGAAAATGAAAGCCAAGTTTCAACAGATGAAAGTGAGAACTCCAGGTCTCCTGGAAATAAA
TCAGATAACATCAAGCCCAAATCTGCTCCATGGAACTCTTTTCTCCCTCCACCACCCCCCATGCCAGGGC
CAAGACTGGGACCAGGAAAGATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGA
TGCTTTGGGAAGTATGTTAATTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGGGTTTTAGA
CAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATTAAGGAGAAATGCTGGCATAGAGCAGCACTA
AATGACACCACTAAAGAAACGATCAGACAGATCTGGAATGTGAAGCGTTATAGAAGATAACTGGCCTCAT
TTCTTCAAAATATCAAGTGTTGGGAAAGAAAAAAGGAAGTGGAATGGGTAACTCTTCTTGATTAAAAGTT
ATGTAATAACCAAATGCAATGTGAAATATTTTACTGGACTCTTTTGAAAAACCATCTGTAAAAGACTGAG
GTGGGGGTGGGAGGCCAGCACGGTGGTGAGGCAGTTGAGAAAATTTGAATGTGGATTAGATTTTGAATGA
TATTGGATAATTATTGGTAATTTTATGGCCTGTGAGAAGGGTGTTGTAGTTTATAAAAGACTGTCTTAAT
TTGCATACTTAAGCATTTAGGAATGAAGTGTTAGAGTGTCTTAAAATGTTTCAAATGGTTTAACAAAATG
TATGTGAGGCGTATGTGGCAAAATGTTACAGAATCTAACTGGTGGACATGGCTGTTCATTGTACTGTTTT
TTTCTATCTTCTATATGTTTAAAAGTATATAATAAAAAATATTTAATTTTTTTTTAAA

Human SMN2 mRNA sequence - var4 (public gi: 13259526) (SEQ ID NO: 150)
CCACAAATGTGGGAGGGCGATAACCACTCGTAGAAAGCGTGAGAAGTTACTACAAGCGGTCCTCCCGGCC
ACCGTACTGTTCCGCTCCCAGAAGCCCCGGGCGGCGGAAGTCGTCACTCTTAAGAAGGGACGGGGCCCCA
CGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCCCGGAGCAGGA
GGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTCTGACATTTGGGATGATACAGCACTG
ATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTAAAGAATGGTGACATTTGTGAAACTT
CGGGTAAACCAAAAACCACACCTAAAAGAAAACCTGCTAAGAAGAATAAAAGCCAAAAGAAGAATACTGC
AGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCATTTGGTCAGAAGACGGTTGCATTTAC
CCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTTGTGGTTTACACTGGATATGGAAATA
GAGAGGAGCAAAATCTGTCCGATCTACTTTCCCCAATCTGTGAAGTAGCTAATAATATAGAACAGAATGC
TCAAGAGAATGAAAATGAAAGCCAAGTTTCAACAGATGAAAGTGAGAACTCCAGGTCTCCTGGAAATAAA
TCAGATAACATCAAGCCCAAATCTGCTCCATGGAACTCTTTTCTCCCTCCACCACCCCCCATGCCAGGGC
CAAGACTGGGACCAGGAAAGCCAGGTCTAAAATTCAATGGCCCACCACCGCCACCGCCACCACCACCACC
CCACTTACTATCATGCTGGCTGCCTCCATTTCCTTCTGGACCACCAATAATTCCCCCACCACCTCCCATA
TGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGTATGTTAATTTCATGGTACATGAGTGGCTATC
ATACTGGCTATTATATGGAAATGCTGGCATAGAGCAGCACTAAATGACACCACTAAAGAAACGATCAGAC
AGATCTGGAATGTGAAGCGTTATAGAAGATAACTGGCCTCATTTCTTCAAAATATCAAGTGTTGGGAAAG
AAAAAAGGAAGTGGAATGGGTAACTCTTCTTGATTAAAAGTTATGTAATAACCAAATGCAATGTGAAATA
TTTTACTGGACTCTTTTGAAAAACCATCTGTAAAAGACTGAGGTGGGGGTGGGAGGCCAGCACGGTGGTG
AGGCAGTTGAGAAAATTTGAATGTGGATTAGATTTTGAATGATATTGGATAATTATTGGTAATTTTATGG
CCTGTGAGAAGGGTGTTGTAGTTTATAAAAGACTGTCTTAATTTGCATACTTAAGCATTTAGGAATGAAG
TGTTAGAGTGTCTTAAAATGTTTCAAATGGTTTAACAAAATGTATGTGAGGCGTATGTGGCAAAATGTTA
CAGAATCTAACTGGTGGACATGGCTGTTCATTGTACTGTTTTTTTCTATCTTCTATATGTTTAAAAGTAT
ATAATAAAAAATATTTAATTTTTTTTTAAA

Human SMN2 mRNA sequence - var5 (public gi: 13259525) (SEQ ID NO: 151)
CCACAAATGTGGGAGGGCGATAACCACTCGTAGAAAGCGTGAGAAGTTACTACAAGCGGTCCTCCCGGCC
ACCGTACTGTTCCGCTCCCAGAAGCCCCGGGCGGCGGAAGTCGTCACTCTTAAGAAGGGACGGGGCCCCA
CGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGCGGCAGTGGTGGCGGCGTCCCGGAGCAGGA
GGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTCTGACATTTGGGATGATACAGCACTG
ATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTAAAGAATGGTGACATTTGTGAAACTT

CGGGTAAACCAAAAACCACACCTAAAAGAAAACCTGCTAAGAAGAATAAAAGCCAAAAGAAGAATACTGC
AGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCATTTGGTCAGAAGACGGTTGCATTTAC
CCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTTGTGGTTTACACTGGATATGGAAATA
GAGAGGAGCAAAATCTGTCCGATCTACTTTCCCCAATCTGTGAAGTAGCTAATAATATAGAACAGAATGC
TCAAGAGAATGAAAATGAAAGCCAAGTTTCAACAGATGAAAGTGAGAACTCCAGGTCTCCTGGAAATAAA
TCAGATAACATCAAGCCCAAATCTGCTCCATGGAACTCTTTTCTCCCTCCACCACCCCCCATGCCAGGGC
CAAGACTGGGACCAGGAAAGCCAGGTCTAAAATTCAATGGCCCACCACCGCCACCGCCACCACCACCACC
CCACTTACTATCATGCTGGCTGCCTCCATTTCCTTCTGGACCACCAATAATTCCCCCACCACCTCCCATA
TGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGTATGTTAATTTCATGGTACATGAGTGGCTATC
ATACTGGCTATTATATGGGTTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATTAAGG
AGAAATGCTGGCATAGAGCAGCACTAAATGACACCACTAAAGAAACGATCAGACAGATCTGGAATGTGAA
GCGTTATAGAAGATAACTGGCCTCATTTCTTCAAAATATCAAGTGTTGGGAAAGAAAAAAGGAAGTGGAA
TGGGTAACTCTTCTTGATTAAAAGTTATGTAATAACCAAATGCAATGTGAAATATTTTACTGGACTCTTT
TGAAAAACCATCTGTAAAAGACTGAGGTGGGGGTGGGAGGCCAGCACGGTGGTGAGGCAGTTGAGAAAAT
TTGAATGTGGATTAGATTTTGAATGATATTGGATAATTATTGGTAATTTTATGGCCTGTGAGAAGGGTGT
TGTAGTTTATAAAAGACTGTCTTAATTTGCATACTTAAGCATTTAGGAATGAAGTGTTAGAGTGTCTTAA
AATGTTTCAAATGGTTTAACAAAATGTATGTGAGGCGTATGTGGCAAAATGTTACAGAATCTAACTGGTG
GACATGGCTGTTCATTGTACTGTTTTTTTCTATCTTCTATATGTTTAAAAGTATATAATAAAAATATTTA
ATTTTTTTTAAA

Human SMN2 Protein sequence - var1 (public gi: 736411) (SEQ ID NO: 276)
AMSSGGSGGGVPEQEDSVLFRRGTGQSDDSDIWDDTALIKAYDKAVASFKHALKNGDICETSGKPKTTPK
RKPAKKNKSQKKNTAASLQQWKVGDKCSAIWSEDGCIYPATIASIDFKRETCVVVYTGYGNREEQNLSDL
LSPICEVANNIEQNAQENENESQVSTDESENSRSPGNKSDNIKPKSAPWNSFLPPPPPMPGPRLGPGKPG
LKFNGPPPPPPPPPPPHLLSCWLPPFPSGPPIIPPPPPICPDSLDDADALGSMLISWYMSGYHTGYYMGFR
QNQKEGRCSHSLN

Human SMN2 Protein sequence - var2 (public gi: 13259531) (SEQ ID NO: 277)
MAMSSGGSGGGVPEQEDSVLFRRGTGQSDDSDIWDDTALIKAYDKAVASFKHALKNGDICETSGKPKTTP
KRKPAKKNKSQKKNTAASLQQWKVGDKCSAIWSEDGCIYPATIASIDFKRETCVVVYTGYGNREEQNLSD
LLSPICEVANNIEQNAQENENESQVSTDESENSRSPGNKSDNIKPKSAPWNSFLPPPPPMPGPRLGPGKI
IPPPPPICPDSLDDADALGSMLISWYMSGYHTGYYMEMLA

Human SMN2 Protein sequence - var3 (public gi: 13259529) (SEQ ID NO: 278)
MAMSSGGSGGGVPEQEDSVLFRRGTGQSDDSDIWDDTALIKAYDKAVASFKHALKNGDICETSGKPKTTP
KRKPAKKNKSQKKNTAASLQQWKVGDKCSAIWSEDGCIYPATIASIDFKRETCVVVYTGYGNREEQNLSD
LLSPICEVANNIEQNAQENENESQVSTDESENSRSPGNKSDNIKPKSAPWNSFLPPPPPMPGPRLGPGKI
IPPPPPICPDSLDDADALGSMLISWYMSGYHTGYYMGFRQNQKEGRCSHSLN

Human SMN2 Protein sequence - var4 (public gi: 13259527) (SEQ ID NO: 279)
MAMSSGGSGGGVPEQEDSVLFRRGTGQSDDSDIWDDTALIKAYDKAVASFKHALKNGDICETSGKPKTTP
KRKPAKKNKSQKKNTAASLQQWKVGDKCSAIWSEDGCIYPATIASIDFKRETCVVVYTGYGNREEQNLSD
LLSPICEVANNIEQNAQENENESQVSTDESENSRSPGNKSDNIKPKSAPWNSFLPPPPPMPGPRLGPGKP
GLKFNGPPPPPPPPPPPHLLSCWLPPFPSGPPIIPPPPPICPDSLDDADALGSMLISWYMSGYHTGYYMEM
LA

Human SMN2 Protein sequence - var5 (public gi: 10937869) (SEQ ID NO: 280)
MAMSSGGSGGGVPEQEDSVLFRRGTGQSDDSDIWDDTALIKAYDKAVASFKHALKNGDICETSGKPKTTP
KRKPAKKNKSQKKNTAASLQQWKVGDKCSAIWSEDGCIYPATIASIDFKRETCVVVYTGYGNREEQNLSD
LLSPICEVANNIEQNAQENENESQVSTDESENSRSPGNKSDNIKPKSAPWNSFLPPPPPMPGPRLGPGKP
GLKFNGPPPPPPPPPPPHLLSCWLPPFPSGPPIIPPPPPICPDSLDDADALGSMLISWYMSGYHTGYYMGF
RQNQKEGRCSHSLN

Unigene Name: SNX1 **Unigene ID:** Hs.498154

Human SNX1 mRNA sequence - var1 (public gi: 3152939) (SEQ ID NO: 152)
ATGGCGTCGGGTGGTGGTGGCTGTAGCGCTTCGGAGAGACTGCCTCCGCCCTTCCCCGGCCTGGAGCCGG
AGTCCGAGGGGGCGGCCGGGGGATCAGAACCCGAGGCTGGGGACAGCGACACCGAGGGGGAGGACATTTT
CACCGGCGCCGCGGTGGTCAGTAAACATCAGTCTCCAAAGATAACTACATCCCTTCTTCCCATCAACAAT

GGCTCCAAAGAAAATGGGATCCATGAAGAACAAGACCAAGAGCCACAGGATCTCTTTGCAGATGCCACAG
TGGAGCTATCCTTGGACAGCACACAAAATAATCAGAAGAAGGTGCTAGCCAAAACACTCATTTCTCTTTC
TCCTCAGGAAGCCACAAATTCTTCGAAGCCCCAGCCAACCTATGAGGAGCTAGAGGAAGAAGAACAGGAG
GATCAATTTGATTTGACAGTCGGTATAACTGATCCTGAGAAGATAGGGGATGGTATGAATGCATATGTAG
CCTACAAAGTTACAACACAGACAAGCTTACCATTGTTCAGAAGCAAACAGTTTGCAGTAAAAAGAAGATT
TAGTGACTTTCTGGGTCTTTATGAGAAGCTTTCCGAGAAGCACTCTCAGAATGGCTTCATTGTCCCTCCA
TCCCCGGAGAAGAGCCTCATAGGGATGACAAAAGTGAAAGTTGGGAAGGAAGATTCTTCTTCTGCAGAAT
TTCTTGAAAAACGGAGGGCCGCTTTAGAAAGGTACCTTCAGAGGATTGTAAATCATCCTACCATGTTACA
GGACCCTGACGTCAGAGAGTTCTTGGAAAAAGAAGAGCTGCCACGTGCCGTGGGTACCCAGACATTGAGT
GGTGCTGGTCTCCTCAAGATGTTCAACAAAGCCACAGATGCCGTCAGCAAAATGACCATCAAGATGAATG
AATCAGACATTTGGTTTGAGGAGAAGCTCCAGGAGGTAGAGTGTGAGGAGCAGCGCTTACGGAAACTGCA
TGCTGTTGTAGAAACTCTAGTCAACCATAGGAAAGAGCTAGCGCTGAACACAGCCCAGTTTGCAAAGAGT
CTAGCCATGCTTGGGAGCTCTGAGGACAACACGGCATTGTCACGGGCACTCTCCCAGCTGGCTGAGGTGG
AAGAAAAAATTGAGCAGCTCCACCAGGAACAGGCCAACAATGACTTCTTCCTCCTTGCTGAGCTCCTGAG
TGACTACATTCGCCTCCTGGCCATAGTCCGCGCTGCCTTCGACCAGCGCATGAAGACATGGCAGCGCTGG
CAGGATGCCCAAGCCACACTGCAGAAGAAGCGGGAGGCCGAGGCTCGGCTGCTGTGGGCCAACAAGCCTG
ATAAGCTGCAGCAGGCCAAGGACGAGATCCTCGAGTGGGAGTCTCGGGTGACTCAATATGAAAGGGACTT
CGAGAGGATTTCAACAGTGGTCCGAAAAGAAGTGATACGGTTTGAGAAAGAGAAATCCAAGGACTTCAAG
AACCACGTGATCAAGTACCTTGAGACACTCCTTTACTCACAGCAGCAGCTGGCAAAGTACTGGGAAGCCT
TCCTTCCTGAGGCAAAGGCCATCTCCTAA

Human SNX1 mRNA sequence - var2 (public gi: 3152941) (SEQ ID NO: 153)
ATGGCGTCGGGTGGTGGTGGCTGTAGCGCTTCGGAGAGACTGCCTCCGCCCTTCCCCGGCCTGGAGCCGG
AGTCCGAGGGGGCGGCCGGGGGATCAGAACCCGAGGCTGGGGACAGCGACACCGAGGGGGAGGACATTTT
CACCGGCGCCGCGGTGGTCAGTAAACATCAGTCTCCAAAGATAACTACATCCCTTCTTCCCATCAACAAT
GGCTCCAAAGAAAATGGGATCCATGAAGAACAAGACCAAGAGCCACAGGATCTCTTTGCAGGGGATGGTA
TGAATGCATATGTAGCCTACAAAGTTACAACACAGACAAGCTTACCATTGTTCAGAAGCAAACAGTTTGC
AGTAAAAAGAAGATTTAGTGACTTTCTGGGTCTTTATGAGAAGCTTTCCGAGAAGCACTCTCAGAATGGC
TTCATTGTCCCTCCATCCCCGGAGAAGAGCCTCATAGGGATGACAAAAGTGAAAGTTGGGAAGGAAGATT
CTTCTTCTGCAGAATTTCTTGAAAAACGGAGGGCCGCTTTAGAAAGGTACCTTCAGAGGATTGTAAATCA
TCCTACCATGTTACAGGACCCTGACGTCAGAGAGTTCTTGGAAAAAGAAGAGCTGCCACGTGCCGTGGGT
ACCCAGACATTGAGTGGTGCTGGTCTCCTCAAGATGTTCAACAAAGCCACAGATGCCGTCAGCAAAATGA
CCATCAAGATGAATGAATCAGACATTTGGTTTGAGGAGAAGCTCCAGGAGGTAGAGTGTGAGGAGCAGCG
CTTACGGAAACTGCATGCTGTTGTAGAAACTCTAGTCAACCATAGGAAAGAGCTAGCGCTGAACACAGCC
CAGTTTGCAAAGAGTCTAGCCATGCTTGGGAGCTCTGAGGACAACACGGCATTGTCACGGGCACTCTCCC
AGCTGGCTGAGGTGGAAGAAAAAATTGAGCAGCTCCACCAGGAACAGGCCAACAATGACTTCTTCCTCCT
TGCTGAGCTCCTGAGTGACTACATTCGCCTCCTGGCCATAGTCCGCGCTGCCTTCGACCAGCGCATGAAG
ACATGGCAGCGCTGGCAGGATGCCCAAGCCACACTGCAGAAGAAGCGGGAGGCCGAGGCTCGGCTGCTGT
GGGCCAACAAGCCTGATAAGCTGCAGCAGGCCAAGGACGAGATCCTCGAGTGGGAGTCTCGGGTGACTCA
ATATGAAAGGGACTTCGAGAGGATTTCAACAGTGGTCCGAAAAGAAGTGATACGGTTTGAGAAAGAGAAA
TCCAAGGACTTCAAGAACCACGTGATCAAGTACCTTGAGACACTCCTTTACTCACAGCAGCAGCTGGCAA
AGTACTGGGAAGCCTTCCTTCCTGAGGCAAAGGCCATCTCCTAA

Human SNX1 mRNA sequence - var3 (public gi: 30582804) (SEQ ID NO: 154)
ATGGCGTCGGGTGGTGGTGGCTGTAGCGCTTCGGAGAGACTGCCTCCGCCCTTCCCCGGCCTGGAGCCGG
AGTCCGAGGGGGCGGCCGGGGGATCAGAACCCGAGGCTGGGGACAGCGACACCGAGGGGGAGGACATTTT
CACCGGCGCCGCGGTGGTCAGTAAACATCAGTCTCCAAAGATAACTACATCCCTTCTTCCCATCAACAAT
GGCTCCAAAGAAAATGGGATCCATGAAGAACAAGACCAAGAGCCACAGGATCTCTTTGCAGATGCCACAG
TGGAGCTATCCTTGGACAGCACACAAAATAATCAGAAGAAGGTGCTAGCCAAAACACTCATTTCTCTTCC
TCCTCAGGAAGCCACAAATTCTTCGAAGCCCCAGCCAACCTATGAGGAGCTAGAGGAAGAAGAACAGGAG
GATCAATTTGATTTGACAGTCGGTATAACTGATCCTGAGAAGATAGGGGATGGTATGAATGCATATGTAG
CCTACAAAGTTACAACACAGACAAGCTTACCATTGTTCAGAAGCAAACAGTTTGCAGTAAAAAGAAGATT
TAGTGACTTTCTGGGTCTTTATGAGAAGCTTTCCGAGAAGCACTCTCAGAATGGCTTCATTGTCCCTCCA
CCCCCGGAGAAGAGCCTCATAGGGATGACAAAAGTGAAAGTTGGGAAGGAAGATTCTTCTTCTGCAGAAT
TTCTTGAAAAACGGAGGGCCGCTTTAGAAAGGTACCTTCAGAGGATTGTAAATCATCCTACCATGTTACA
GGACCCTGACGTCAGAGAGTTCTTGGAAAAAGAAGAGCTGCCACGTGCCGTGGGTACCCAGACATTGAGT
GGTGCTGGTCTCCTCAAGATGTTCAACAAAGCCACAGATGCCGTCAGCAAAATGACCATCAAGATGAATG
AATCAGACATTTGGTTTGAGGAGAAGCTCCAGGAGGTAGAGTGTGAGGAGCAGCGCTTACGGAAACTGCA
TGCTGTTGTAGAAACTCTAGTCAACCATAGGAAAGAGCTAGCGCTGAACACAGCCCAGTTTGCAAAGAGT
CTAGCCATGCTTGGGAGCTCTGAGGACAACACGGCATTGTCACGGGCACTCTCCCAGCTGGCTGAGGTGG
AAGAAAAAATTGAGCAGCTCCACCAGGAACAGGCCAACAATGACTTCTTCCTCCTTGCTGAGCTCCTGAG
TGACTACATTCGCCTCCTGGCCATAGTCCGCGCTGCCTTCGACCAGCGCATGAAGACATGGCAGCGCTGG

CAGGATGCCCAAGCCACACTGCAGAAGAAGCGGGAGGCCGAGGCTCGGCTGCTGTGGGCCAACAAGCCTG
ATAAGCTGCAGCAGGCCAAGGACGAGATCCTCGAGTGGGAGTCTCGGGTGACTCAATATGAAAGGGACTT
CGAGAGGATTTCAACAGTGGTCCGAAAAGAAGTGATACGGTTTGAGAAAGAGAAATCCAAGGACTTCAAG
AACCACGTGATCAAGTACCTTGAGACACTCCTTTACTCACAGCAGCAGCTGGCAAAGTACTGGGAAGCCT
TCCTTCCTGAGGCAAAGGCCATCTCCTAG

Human SNX1 mRNA sequence - var4 (public gi: 4884359) (SEQ ID NO: 155)
GGTTGCTTTGTTAAGTTCCATCTAATGATCATTCTGACGTAAGTCTGTTTTTCTTATTTCCTTGGAATGA
TGTCTCCTCTGGTTTCAGAACTTCCTCCTCTGCTTCCTGTATCCTGAGGCTGGCGGGGCCAGTTGTCTTT
AGGGCTTGTGCATTTTTGTAAAGAGCTTGCACGTGTGGAAATCAAGTAGGCCAGTAGTGGGTTAGGGGTA
CTGAGCCAGAAGCCTCTACAAGGAATAACAGGAGCACAAAGGAAGAAGGTGGTATTCCAGCTGGGGACCC
AGGAGGGAGGACTTTGTGGAGAACCTGATGCTTGAACTGAGTCTAAAAGGTGTAAAAGTGTTTGTTGCTT
CTGCCTCCCTGTCTGTCTGGCAGGGTGAGGTAGGCGCATCTAGGGAAATGTCAAGTGGCTTGGTGTAGGG
TAAAGTCAGTGAGGCCCATGGAGAAAAACGAGCAGGAGCCACATCACATGGGTGTCTGATAGGACCTGGG
AGGCGCTTTCCACATTACCATTGTCGCTTCGTGATCTGGACACACCAGAAGGCGTGAGACTGGAGGCAGG
AAGAGCAGCCAGGCTTATCCCTACCCTCAGGAGAGCTGAAAAGGGCAGGTATGGTGGGGCCAGAGCTCAG
GAGAGTTTCGGAACCACTGAGATCGGTCCTTGATTTGATGAGAGGCTTGAGGGGAGAGGGAGGTAGCTAG
GATGCCCCGCAAGCTTCTGGCCCAGACACTGGGCAGACAATGAAACCCTTTGTAACACATGAGGCAATAG
GTTTGGGGCAGATGGGAGGGGAAGCAGTGGTGGGGGCAGTGAGTGCTGAAGGTGTTTTAAGAAGCGGCTC
TGGGCCAGGCACAGTGGCTTATGCCTGTATTCCTAGCATTTTGGGAGGCCGAGGTGGGAGAATCACTTGA
GCCCAAGAATTTGAGACCAGCCTGGGGAATATAGTGAGACCCTGTCCCTACAAAAATAAAAACAACTAGC
TGGGTGTGTGGTGGTGCATGCCTGTAGGCCCAGCTACGCGGGAACATCACCTGAGCCCAGGAGGTTGAGG
TTGCAGTGAGCTACAGTTGCGCCACTGCACTCCAGCCTAGGTGACAGAGCAAGATCTTGTCTCAAAAAAA
AAAACAGCTCTGGATGGGAAGGGAGGCCAGTTGCTTTAAGTAGGGGAGATAGAGTTAAAGGAGGCTTTGT
TTTATTTAAAGGTGGGACAAACTTAAGCATGTTAATAAAATTCAGAGAAGAGAAAGAGAATGACTATCAG
AGCCATGTTTGGAAGAAAATGGGGTCCAGAGCACAGGAAGGGGACCTGTGTTCAGAGGGTGCCTCACTGC
TGAGGCCACAGGAAAGAATCTGTAGGTGGAGGGGAGGCCGAAGAGGGGAAGTTTCATGCTTGATAATTAA
AATTTTCTGAGATAGGAATGTCATATTTACCTATTTAAGCCAAGTTTTTTAGATAAAAGGTATGGAACC
TGCTTTCCCCTTGGCTAGTTCAGCGTTTGGGCTCCGGAGTGCTGAAGATGAGGACTGGACTTCGAGCTGG
TGTGATCCCAGTATTCAGTGTCAGTACTCAGTGACAAAATAAATGAGAGAAACGGGAATAAGAATTGTCG
CCTACACAAAAATACCAGCAACTGTTAACTCTTCCCAGAAGATTTTCATTCTGAATGCTCCTGTAGCTAG
GAACCCTAAAAAGTCTTTGAAGCAACTCAAGTTTTAAAAAAGGGGAGGAACTCCTGGAAATCTCAGGATG
GGGCCAAGATGTGGCTGGAGAGTGTGTGGTGATGGAGGGCGTGTCTTTTGCCGAGCACACTCAGGGCCCA
CGGGAAGCCCATAGACTTCAAGGACATCAAGCCCCAAGGTGGTGGGATTTTCCCCACCAGTACTTGGCAG
CCTAGGGGGAAGGGGAGGGCGGGAGAAGATAATGGGGATCCCTGGCTCCAAACATAGGAGGACACATCTG
TGCTACAGTGCGCACATGCCTGGATGTACACTCTGTCTTTGGAGACACTGGCTAAGATTCTCTGCTCCAT
GTTTGGACAGGGTCGTGCCTGATCTGAGATAAATGGACAAGAACAACTGAAGCCTGTCTTCTGGTGCATG
TGTCACCTGCCGATAACTGCATCTTGTGATAAAGTTGGGTGATTTACAGTCTCCACCAAATGCTAAACTC
TGGGGTCTTACGCCTTTATAACTCCATGGGCCCCAGCAAAGGTTCAGGCTCAAAACAGGTGTCAAATAGA
TAACTGTTGAATGATTGTTCCCCAGTTGCAGGCTCTGCCACCTGGCGTTCATACTGTCTGTGAAAGGACC
CAGCTCACCTTTCCCTCTTTATCTCCCAGTCCTTCCCAACAGCGCCGACACCTCATGGAAACTGATTGCA
AATGTGCTACTTCTCACTTCTGTGTGGCCCGAGGAGGCTGGGTTAATGCTGGGCTTGGTACCTTAAGCAC
CCTTTCTCCCTTCCCCATCTTCATTCTCAGAATTACACCTGTCTGAAGCAGGCATTTTCCAATGCCCTAG
ATGGGAATATAAGTGTAAGGAGATGTGAAGCATTTGCCTGTGTGTCAGAACATTCACTGAGGATCCTCAT
AGGCACTTCTAGAAACCAAATCCTTGAAGATGACTAACCAGAAATGCCCGTCATAGCACTGTTTACAGTT
GCAAAAACTGAAGCCAATTGAAATGTCCATCAGGAGGGGATTAAATGAATTATGGTACAGTTACACCGTT
GAATATTTTACAGCCATTGAAGATGATATATAGCTATATTCATTGACAAGGAAAACTCATATTTTTTAGT
GAAAAAAGCAGGTTATAGAATTGCATGATATTCACATTTATATAAAACTTTATATATGGGAAGGATGTTG
ATTGAATTGTTAATAACTATGGTCACCTCTAGAGATGGAAGTTTGCATTACCTTTAATTTTTAATACCAT
TTTGTATTGCTTAAAATTTGTATGTATTATCGTTAAAATAAGAAAAATCAAATAAAGCTATTTTCATTAT
GGGAAAAAAAAAAAAAAAAA

Human SNX1 mRNA sequence - var5 (public gi: 4406620) (SEQ ID NO: 156)
ATAAAAGGTATGGAACCTGCTTTCCCCTTGGCTAGTTCAGCGTTTGGGCTCCGGAGTGCTGAAGATGAGG
ACTGGACTTCGAGCTGGTGTGATCCCAGTATTCAGTGTCAGTACTCAGTGACAAAATAAATGAGAGAAAC
GGGAATAAGAATTGTCGCCTACACAAAAATACCAGCAACTGTTAACTCTTCCCAGAAGATTTTCATTCTG
AATGCTCCTGTAGCTAGGAACCCTAAAAAGTCTTTGAAGCAACTCAAGTTTTAAAAAAGGGGAGGAACTC
CTGGAAATCTCAGGATGGGGCCAAGATGTGGCTGGAGAGTGTGTGGTGATGGAGGGCGTGTCTTTTGCCG
AGCACACTCAGGGCCCACGGGAAGCCCATAGACTTCAAGGACATCAAGCCCCAAGGTGGTGGGATTTTCC
CCACCAGTACTTGGCAGCCTAGGGGGAAGGGGAGGGCGGGAGAAGATAATGGGGATCCCTGGCTCCAAAC
ATAGGAGGACACATCTGTGCTACAGTGCGCACATGCCTGGATGTACACTCTGTCTTTGGAGACACTGGCT
AAGATTCTCTGCTCCATGTTTGGACAGGGTCGTGCCTGATCTGAGATAAATGGACAAGAACAACTGAAGC

CTGTCTTCTGGTGCATGTGTCACCTGCCGATAACTGCATCTTGTGATAAAGTTGGGTGATTTACAGTCTC
CACCAAATGCTAAACTCTGGGGTCTTACGCCTTTATAACTCCATGGGCCCCAGCAAAGGTTCAGGCTCAA
AACAGGTGTCAAATAGATAACTGTTGAATGATTGTTCCCCAGTTGCAGGCTCTGCCACCTGGCGTTCATA
CTGTCTGTGAAAGGACCCAGCTCACCTTTCCCTCTTTATCTCCCAGTCCTTCCCAACAGCGCCGACACCT
CATGGAAACTGATTGCAAATGTGCTACTTCTCACTTCTGTGTGGCCCGAGGAGGCTGGGTTAATGCTGGG
CTTGGTACCTTAAGCACCCTTTCTCCCTTCCCCATCTTCATTCTCAGAATTACACCTGTCTGAAGCAGGC
ATTTTCCAATGCCCTAGATGGGAATATAAGTGTAAGGAGATGTGAAGCATTTGCCTGTGTGTCAGAACAT
TCACTGAGGATCCTCATAGGCACTTCTAGAAACCAAATCCTTGAAGATGACTAACCAGAAATGCCCGTCA
TAGCACTGTTTACAGTTGCAAAAACTGAAGCCAATTGAAATGTCCATCAGGAGGGGATTAAATGAATTAT
GGTACAGTTACACCGTTGAATATTTTACAGCCATTGAAGATGATATATAGCTATATTCATTGACAAGGAA
AACTCATATTTTTTAGTGAAAAAAGCAGGTTATAGAATTGCATGATATTCACATTTATATAAAACTTTAT
ATATGGGAAGGATGTTGATTGAATTGTTAATAACTATGGTCACCTCTAGAGATGGAAGTTTGCATTACCT
TTAATTTTTAATACCATTTTGTATTGCTTAAAATTTGTATGTATTATCGTTAAAATAAGAAAAATCAAAT
AAAGCTATTTTCATTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human SNX1 mRNA sequence - var6 (public gi: 34535422) (SEQ ID NO: 157)
TTTCCGCCGCGGGTGGAAGAAGATGGCGTCGGGTGGTGGTGGCTGTAGCGCTTCGGAGAGACTGCCTCCG
CCCTTCCCCGGCCTGGAGCCGGAGTCCGAGGGGGCGGCCGGGGGATCAGAACCCGAGGCTGGGGACAGCG
ACACCGAGGGGGAGGACATTTTCACCGGCGCCGCGGTGGTCAGTAAACATCAGTCTCCAAAGATAACTAC
ATCCCTTCTTCCCATCAACAATGGCTCCAAAGAAAATGGATCCATGAAGAACAAGACCAAGAGCCACAG
GATCTCTTTGCAGATGCCACAGTGGAGCTATCCTTGGACAGCACACAAAATAATCAGAAGAAGGTGCTAG
CCAAAACACTCATTTCTCTTCCTCCTCAGGAAGCCACAAATTCTTCGAAGCCCCAGCCAACCTATGAGGA
GCTAGAGGAAGAAGAACAGGAGGATCAATTTGATTTGACAGTCGGTATAACTGATCCTGAGAAGATAGGG
GATGGTATGAATGCATATGTAGCCTACAAAGTTACAACACAGACAAGCTTACCATTGTTCAGAAGCAAAC
AGTTTGCAGTAAAAAGAAGATTTAGTGACTTTCTGGGTCTTTATGAGAAGCTTTCCGAGAAGCACTCTCA
GAATGGCTTCATTGTCCCTCCACCCCCGGAGAAGAGCCTCATAGGGATGACAAAAGTGAAAGTTGGGAAG
GAAGATTCTTCTTCTGCAGAATTTCTTGAAAAACGGAGGGCCGCTTTAGAAAGGTACCTTCAGAGGATTG
TAAATCATCCTACCATGTTACAGGACCCTGACGTCAGAGAGTTCTTGGAAAAAGAAGAGCTGCCACGTGC
CGTGGGTACCCAGACATTGAGTGGTGCTGGTCTCCTCAAGATGTTCAACAAAGCCACAGATGCCGTCAGC
AAAATGACCATCAAGATGAATGAATCAGACATTTGGTTTGAGGAGAAGCTCCAGGAGGTAGAGTGTGAGG
AGCAGCGCTTACGGAAACTGCATGCTGTTGTAGAAACTCTAGTCAACCATAGGAAAGAGCTAGCGCTGAA
CACAGCCCAGTTTGCAAAGAGTCTAGCCATGCTTGGGAGCTCTGAGGACAACACGGCATTGTCACGGGCA
CTCTCCCAGCTGGCTGAGGTGGAAGAAAAAATTGAGCAGCTCCACCAGGAACAGGCCAACAATGACTTCT
TCCTCCTTGCTGAGCTCCTGAGTGACTACATTCGCCTCCTGGCCATAGTCCGCGCTGCCTTCGACCAGCG
CATGAAGACATGGCAGCGCTGGCAGGATGCCCAAGCCACACTGCAGAAGAAGCGGGAGGCCGAGGCTCGG
CTGCTGTGGGCCAACAAGCCTGATAAGCTGCAGCAGGCCAAGGACGAGATCCTCGAGTGGGAGTCTCGGG
TGACTCAATATGAAAGGGACTTCGAGAGGATTTCAACAGTGGTCCGAAAAGAAGTGATACGGTTTGAGAA
AGAGAAATCCAAGGACTTCAAGAACCACGTGATCAAGTACCTTGAGACACTCCTTTGCTCACAGCAGCAG
GCTGGGGAGCAGTTGGGAATCAGGTCTGGAATACTCCTAACCAAGAAGTTGCCCAGGTATAGTAAGTTTT
TCTCTACCGTTCACAAGTTTTGTGCTGCTGCTTCCCTCTGGAAATGGGGTTTCTTTCTCTCCGCCTACCT
CAGCTACCTGTTCTGAGGGTCTCAATCTGTTTCGTATTCCCACTTCTTTAGGGAAGGAGTTTTAAAAACA
TCTCTTAAAATAAGAGGAGCAAAATCTATTAAAACCTATTCTCCTGCAAAGGAGGCAGAGACTTTCTCTC
TCTCTTTTTTTTTTTTTTTTGGTGTCCCTATCATTAAGCAAGAGCCTTTCTCTTTTATTCTTCCTGCTT
CCCTAAGCTGCTCAGGGCTCTCTGAGTCTTGCCCTCTGATGGCAAGTCTTATATATAACTAAACCTATTT
TTGTCACCCATCAAAACACATCCTCAGTAGACTGTGTGAAGGTGTGAAGGTCTGATAATGACTTGATGCT
TTATCTCCATAGACATGAAAGCCATGCCCTCTGCCTCTAGATAGGGTGATCCAAGAGCTCCTGAACCTTA
GGAGGTTCAAAGAAGCTCTACTGTCTGTGCCCAGGAGGTAGCCTGCCAGCAAGAGCCCTCAGGAGTTGCA
CACACAGCCAAAGGGTGTTCACACAGATCTCTGCCGGTCTAGCCAGGGGAGGCCAGAGTCTCGTCAGTCA
AGGATGGGCTTCCCCCTTAGCTGTGTCCACAGCTGCTCAAGCTATACTGGTCAGAGTGGGCTTTGAAGCT
CCTTTGTGAGCTCGAGCTGCTGACTGCCACTATGGGAGCCTTGCCACCTCCAGCCCCTCCATCCCAAAGA
CGCTCCTGCCACTGGGGCCCCAGGTCCTGCTGTATCAGTTCTCTTTGGTGGGGGGCTAAGGTTTGGGGCG
AGGCAACCTGAGACAAGAAAACGCAGTAAACATTCTGATTCCCTGTACACAGATGCAGCACCAGGGGAAG
GGCCAGTGGTGCAAGTATTTCTTTTTAACAGGTGAAGTTTTTGGAAAAAGTCACTCTCCCTACCCCTCAG
TATCCTTACCATCAACTTTGGTTTTATCCTTCCAGTCTTTATTATATGCTTGCTTTTACATAGTTGTAAT
AATATACACATAAAGTATTTTGTATCCTGCTTTTATCATTCAACATTGTACATGTTATAAGCATTTTACT
ATATTGTTATATATCTTCACAAAGTTGATCTGTAAAGCTGTGTAATTTGAAGGCATCCATAGGGTGACTG
TACCATAATTTTGATTCATCCCTTGTTGTTGGATTCTTGGTCAGGGGTTGTTTGTTTTGTTTTATTGGTA
ACTTTAAAATTTTGAATACAATTTCAGATTTACAGAAAAGTTGCAGGAATATCACAAAGAACTCCTATAT
ATCTTTTATCCAGATTTACTGAGTGTTTACATTTTATCCCATTTGCCTTATCTATATTTCATGTTGCATT
TTCTTAATCATTTGAGAATAATTTGCACAGATACCCCATTATGCCCAAAACAGTATGCATTTCCCTAAGA
ACAGGACATTCTCTTCTAAGAGAAGAAGAGAATTACTTTAAGCATTATTCAGTATTTTTTTAAGTATTAT
TATCAAAATCAGGAAGTTTAACAGTGATTTAATACTGTTATCTAACCCATGATTCATATTTAAATTTTGC
CATTTATCCCAATAATGTCCTTTGTAGCCATTCTTTTACCTTGTGCAGGATCATGTTACATTTGTAAACG

TGTGTCTCTCAATACTGCAGATTCCTCAACTTTCTTTTGTCTTTCATTACCATGACATTTTTGAAGAATA
CAGGCTATTTTGTCG

Human SNX1 mRNA sequence - var7 (public gi: 38197125) (SEQ ID NO: 158)
GTGGAAGAAGATGGCGTCGGGTGGTGGTGGCTGTAGCGCTTCGGAGAGACTGCCTCCGCCCTTCCCCGGC
CTGGAGCCGGAGTCCGAGGGGGCGGCCGGGGGATCAGAACCCGAGGCTGGGGACAGCGACACCGAGGGGG
AGGACATTTTCACCGGCGCCGCGGTGGTCAGTAAACATCAGTCTCCAAAGATAACTACATCCCTTCTTCC
CATCAACAATGGCTCCAAAGAAAATGGGATCCATGAAGAACAAGACCAAGAGCCACAGGATCTCTTTGCA
GATGCCACAGTGGAGCTATCCTTGGACAGCACACAAAATAATCAGAAGAAGGTGCTAGCCAAAACACTCA
TTTCTCTTCCTCCTCAGGAAGCCACAAATTCTTCGAAGCCCCAGCCAACCTATGAGGAGCTAGAGGAAGA
AGAACAGGAGGATCAATTTGATTTGACAGTCGGTATAACTGATCCTGAGAAGATAGGGGATGGTATGAAT
GCATATGTAGCCTACAAAGTTACAACACAGACAAGCTTACCATTGTTCAGAAGCAAACAGTTTGCAGTAA
AAAGAAGATTTAGTGACTTTCTGGGTCTTTATGAGAAGCTTTCCGAGAAGCACTCTCAGAATGGCTTCAT
TGTCCCTCCACCCCCGGAGAAGAGCCTCATAGGGATGACAAAAGTGAAAGTTGGGAAGGAAGATTCTTCT
TCTGCAGAATTTCTTGAAAAACGGAGGGCCGCTTTAGAAAGGTACCTTCAGAGGATTGTAAATCATCCTA
CCATGTTACAGGACCCTGACGTCAGAGAGTTCTTGGAAAAAGAAGAGCTGCCACGTGCCGTGGGTACCCA
GACATTGAGTGGTGCTGGTCTCCTCAAGATGTTCAACAAAGCCACAGATGCCGTCAGCAAAATGACCATC
AAGATGAATGAATCAGACATTTGGTTTGAGGAGAAGCTCCAGGAGGTAGAGTGTGAGGAGCAGCGCTTAC
GGAAACTGCATGCTGTTGTAGAAACTCTAGTCAACCATAGGAAAGAGCTAGCGCTGAACACAGCCCAGTT
TGCAAAGAGTCTAGCCATGCTTGGGAGCTCTGAGGACAACACGGCATTGTCACGGGCACTCTCCCAGCTG
GCTGAGGTGGAAGAAAAAATTGAGCAGCTCCACCAGGAACAGGCCAACAATGACTTCTTCCTCCTTGCTG
AGCTCCTGAGTGACTACATTCGCCTCCTGGCCATAGTCCGCGCTGCCTTCGACCAGCGCATGAAGACATG
GCAGCGCTGGCAGGATGCCCAAGCCACACTGCAGAAGAAGCGGGAGGCCGAGGCTCGGCTGCTGTGGGCC
AACAAGCCTGATAAGCTGCAGCAGGCCAAGGACGAGATCCTCGAGTGGGAGTCTCGGGTGACTCAATATG
AAAGGGACTTCGAGAGGATTTCAACAGTGGTCCGAAAAGAAGTGATACGGTTTGAGAAAGAGAAATCCAA
GGACTTCAAGAACCACGTGATCAAGTACCTTGAGACACTCCTTTACTCACAGCAGCAGCTGGCAAAGTAC
TGGGAAGCCTTCCTTCCTGAGGCAAAGGCCATCTCCTAATGGACCAAGGACCCCAGAGCCCACCTGTGTG
ACGCTGCCTTTTTATACACTGTCCTCCTCCACCTTGATGGACCCCTAGTGATGCATCCTGCCTAGGCTGG
ACTTAACCCCTTCCTCCCTGTCCCCACGACCAACTGTCCCCAGTTACTCTAACCGTTATTTCATTTAGCT
TCCATATATATTTCTTACCTAAGAGAATAGTTTCCTGCTTTAAGCAAAAGACCTACAATAGGTGGTGGA
ATTATGGGATGGGGTGGAGTATTGATATAAATATATAAATACAAATGTATATTTTTCAGGATGTGGTTTA
GGAACTGGGAATAACGTTTTCTGTTACTCCTGATGGTGCCATGAAAAGGTTATGTAATAAAATATTTTAA
AATCAAAAAAAAAAAAAAAAAA

Human SNX1 mRNA sequence - var8 (public gi: 23111033) (SEQ ID NO: 159)
GGGTGGAAGAAGATGGCGTCGGGTGGTGGTGGCTGTAGCGCTTCGGAGAGACTGCCTCCGCCCTTCCCCG
GCCTGGAGCCGGAGTCCGAGGGGGCGGCCGGGGGATCAGAACCCGAGGCTGGGGACAGCGACACCGAGGG
GGAGGACATTTTCACCGGCGCCGCGGTGGTCAGTAAACATCAGTCTCCAAAGATAACTACATCCCTTCTT
CCCATCAACAATGGCTCCAAAGAAAATGGGATCCATGAAGAACAAGACCAAGAGCCACAGGATCTCTTTG
CAGATGCCACAGTGGAGCTATCCTTGGACAGCACACAAAATAATCAGAAGAAGGTGCTAGCCAAAACACT
CATTTCTCTTCCTCCTCAGGAAGCCACAAATTCTTCGAAGCCCCAGCCAACCTATGAGGAGCTAGAGGAA
GAAGAACAGGAGGATCAATTTGATTTGACAGTCGGTATAACTGATCCTGAGAAGATAGGGGATGGTATGA
ATGCATATGTAGCCTACAAAGTTACAACACAGACAAGCTTACCATTGTTCAGAAGCAAACAGTTTGCAGT
AAAAAGAAGATTTAGTGACTTTCTGGGTCTTTATGAGAAGCTTTCCGAGAAGCACTCTCAGAATGGCTTC
ATTGTCCCTCCGCCCCCGGAGAAGAGCCTCATAGGGATGACAAAAGTGAAAGTTGGGAAGGAAGATTCTT
CTTCTGCAGAATTTCTTGAAAAACGGAGGGCCGCTTTAGAAAGGTACCTTCAGAGGATTGTAAATCATCC
TACCATGTTACAGGACCCTGACGTCAGAGAGTTCTTGGAAAAAGAAGAGCTGCCACGTGCCGTGGGTACC
CAGACATTGAGTGGTGCTGGTCTCCTCAAGATGTTCAACAAAGCCACAGATGCCGTCAGCAAAATGACCA
TCAAGATGAATGAATCAGACATTTGGTTTGAGGAGAAGCTCCAGGAGGTAGAGTGTGAGGAGCAGCGCTT
ACGGAAACTGCATGCTGTTGTAGAAACTCTAGTCAACCATAGGAAAGAGCTAGCGCTGAACACAGCCCAG
TTTGCAAAGAGTCTAGCCATGCTTGGGAGCTCTGAGGACAACACGGCATTGTCACGGGCACTCTCCCAGC
TGGCTGAGGTGGAAGAAAAAATTGAGCAGCTCCACCAGGAACAGGCCAACAATGACTTCTTCCTCCTTGC
TGAGCTCCTGAGTGACTACATTCGCCTCCTGGCCATAGTCCGCGCTGCCTTCGACCAGCGCATGAAGACA
TGGCAGCGCTGGCAGGATGCCCAAGCCACACTGCAGAAGAAGCGGGAGGCCGAGGCTCGGCTGCTGTGGG
CCAACAAGCCTGATAAGCTGCAGCAGGCCAAGGACGAGATCCTCGAGTGGGAGTCTCGGGTGACTCAATA
TGAAAGGGACTTCGAGAGGATTTCAACAGTGGTCCGAAAAGAAGTGATACGGTTTGAGAAAGAGAAATCC
AAGGACTTCAAGAACCACGTGATCAAGTACCTTGAGACACTCCTTTACTCACAGCAGCAGCTGGCAAAGT
ACTGGGAAGCCTTCCTTCCTGAGGCAAAGGCCATCTCCTAATGGACCAAGGACCCCAGAGCCCACCTGTG
TGACGCTGCCTTTTTATACACTGTCCTCCTCCACCTTGATGGACCCCTAGTGATGCATCCTGCCTAGGCT
GGACTTAACCCCTTCCTCCCTGTCCCCACGACCAACTGTCCCCAGTTACTCTAACCGTTATTTCATTTAG
CTTCCATATATATTTCTTACCTAAGAGAATAGTTTCCTGCTTTAAGCAAAAGACCTACAATAGGTGGTG
GAATTATGGGATGGGGTGGAGTATTGATATAAATATATAAATACAAATGTATATTTTTCAGGATGTGGTT

TAGGAACTGGGAATAACGTTTTCTGTTACTCCTGATGGTGCCATGAAAAGGTTATGTAATAAAATATTTT
AAAATCAAAAAAAAAAAAAAAAAA

Human SNX1 mRNA sequence - var9 (public gi: 23111035) (SEQ ID NO: 160)
GGGTGGAAGAAGATGGCGTCGGGTGGTGGTGGCTGTAGCGCTTCGGAGAGACTGCCTCCGCCCTTCCCCG
GCCTGGAGCCGGAGTCCGAGGGGGCGGCCGGGGGATCAGAACCCGAGGCTGGGGACAGCGACACCGAGGG
GGAGGACATTTTCACCGGCGCCGCGGTGGTCAGTAAACATCAGTCTCCAAAGATAACTACATCCCTTCTT
CCCATCAACAATGGCTCCAAAGAAAATGGGATCCATGAAGAACAAGACCAAGAGCCACAGGATCTCTTTG
CAGATGCCACAGTGGAGCTATCCTTGGACAGCACACAAAATAATCAGAAGAAGGTGCTAGCCAAAACACT
CATTTCTCTTCCTCCTCAGGAAGCCACAAATTCTTCGAAGCCCCAGCCAACCTATGAGGAGCTAGAGGAA
GAAGAACAGGAGGATCAATTTGATTTGACAGTCGGTATAACTGATCCTGAGAAGATAGGGGATGGTATGA
ATGCATATGTAGCCTACAAAGTTACAACACAGACAAGCTTACCATTGTTCAGAAGCAAACAGTTTGCAGT
AAAAAGAAGATTTAGTGACTTTCTGGGTCTTTATGAGAAGCTTTCCGAGAAGCACTCTCAGAATGGCTTC
ATTGTCCCTCCGCCCCCGGAGAAGAGCCTCATAGGGATGACAAAAGTGAAAGTTGGGAAGGAAGATTCTT
CTTCTGCAGAATTTCTTGAAAAACGGAGGGCCGCTTTAGAAAGGTACCTTCAGAGGATTGTAAATCATCC
TACCATGTTACAGGACCCTGACGTCAGAGAGTTCTTGGAAAAAGAAGAGCTGCCACGTGCCGTGGGTACC
CAGACATTGAGTGGTGCTGGTCTCCTCAAGATGTTCAACAAAGCCACAGATGCCGTCAGCAAAATGACCA
TCAAGATGAATGAATCAGACATTTGGTTTGAGGAGAAGCTCCAGGAGGTAGAGTGTGAGGAGCAGCGCTT
ACGGAAACTGCATGCTGTTGTAGAAACTCTAGTCAACCATAGGAAAGAGCTAGCGCTGAACACAGCCCAG
TTTGCAAAGAGTCTAGCCATGCTTGGGAGCTCTGAGGACAACACGGCATTGTCACGGGCACTCTCCCAGC
TGGCTGAGGTGGAAGAAAAAATTGAGCAGCTCCACCAGGAACAGGCCAACAATGACTTCTTCCTCCTTGC
TGAGCTCCTGAGTGACTACATTCGCCTCCTGGCCATAGTCCGCTGGGAGTCTCGGGTGACTCAATATGAA
AGGGACTTCGAGAGGATTTCAACAGTGGTCCGAAAAGAAGTGATACGGTTTGAGAAAGAGAAATCCAAGG
ACTTCAAGAACCACGTGATCAAGTACCTTGAGACACTCCTTTACTCACAGCAGCAGCTGGCAAAGTACTG
GGAAGCCTTCCTTCCTGAGGCAAAGGCCATCTCCTAATGGACCAAGGACCCCAGAGCCCACCTGTGTGAC
GCTGCCTTTTTATACACTGTCCTCCTCCACCTTGATGGACCCCTAGTGATGCATCCTGCCTAGGCTGGAC
TTAACCCCTTCCTCCCTGTCCCCACGACCAACTGTCCCCAGTTACTCTAACCGTTATTTCATTTAGCTTC
CATATATATTTTCTTACCTAAGAGAATAGTTTCCTGCTTTAAGCAAAAGACCTACAATAGGTGGTGGAAT
TATGGGATGGGGTGGAGTATTGATATAAATATATAAATACAAATGTATATTTTTCAGGATGTGGTTTAGG
AACTGGGAATAACGTTTTCTGTTACTCCTGATGGTGCCATGAAAAGGTTATGTAATAAAATATTTTAAAA
TCAAAAAAAAAAAAAAAAAA

Human SNX1 mRNA sequence - var10 (public gi: 23111031) (SEQ ID NO: 161)
GGGTGGAAGAAGATGGCGTCGGGTGGTGGTGGCTGTAGCGCTTCGGAGAGACTGCCTCCGCCCTTCCCCG
GCCTGGAGCCGGAGTCCGAGGGGGCGGCCGGGGGATCAGAACCCGAGGCTGGGGACAGCGACACCGAGGG
GGAGGACATTTTCACCGGCGCCGCGGTGGTCAGTAAACATCAGTCTCCAAAGATAACTACATCCCTTCTT
CCCATCAACAATGGCTCCAAAGAAAATGGGATCCATGAAGAACAAGACCAAGAGCCACAGGATCTCTTTG
CAGGGGATGGTATGAATGCATATGTAGCCTACAAAGTTACAACACAGACAAGCTTACCATTGTTCAGAAG
CAAACAGTTTGCAGTAAAAAGAAGATTTAGTGACTTTCTGGGTCTTTATGAGAAGCTTTCCGAGAAGCAC
TCTCAGAATGGCTTCATTGTCCCTCCGCCCCCGGAGAAGAGCCTCATAGGGATGACAAAAGTGAAAGTTG
GGAAGGAAGATTCTTCTTCTGCAGAATTTCTTGAAAAACGGAGGGCCGCTTTAGAAAGGTACCTTCAGAG
GATTGTAAATCATCCTACCATGTTACAGGACCCTGACGTCAGAGAGTTCTTGGAAAAAGAAGAGCTGCCA
CGTGCCGTGGGTACCCAGACATTGAGTGGTGCTGGTCTCCTCAAGATGTTCAACAAAGCCACAGATGCCG
TCAGCAAAATGACCATCAAGATGAATGAATCAGACATTTGGTTTGAGGAGAAGCTCCAGGAGGTAGAGTG
TGAGGAGCAGCGCTTACGGAAACTGCATGCTGTTGTAGAAACTCTAGTCAACCATAGGAAAGAGCTAGCG
CTGAACACAGCCCAGTTTGCAAAGAGTCTAGCCATGCTTGGGAGCTCTGAGGACAACACGGCATTGTCAC
GGGCACTCTCCCAGCTGGCTGAGGTGGAAGAAAAAATTGAGCAGCTCCACCAGGAACAGGCCAACAATGA
CTTCTTCCTCCTTGCTGAGCTCCTGAGTGACTACATTCGCCTCCTGGCCATAGTCCGCGCTGCCTTCGAC
CAGCGCATGAAGACATGGCAGCGCTGGCAGGATGCCCAAGCCACACTGCAGAAGAAGCGGGAGGCCGAGG
CTCGGCTGCTGTGGGCCAACAAGCCTGATAAGCTGCAGCAGGCCAAGGACGAGATCCTCGAGTGGGAGTC
TCGGGTGACTCAATATGAAAGGGACTTCGAGAGGATTTCAACAGTGGTCCGAAAAGAAGTGATACGGTTT
GAGAAAGAGAAATCCAAGGACTTCAAGAACCACGTGATCAAGTACCTTGAGACACTCCTTTACTCACAGC
AGCAGCTGGCAAAGTACTGGGAAGCCTTCCTTCCTGAGGCAAAGGCCATCTCCTAATGGACCAAGGACCC
CAGAGCCCACCTGTGTGACGCTGCCTTTTTATACACTGTCCTCCTCCACCTTGATGGACCCCTAGTGATG
CATCCTGCCTAGGCTGGACTTAACCCCTTCCTCCCTGTCCCCACGACCAACTGTCCCCAGTTACTCTAAC
CGTTATTTCATTTAGCTTCCATATATATTTTCTTACCTAAGAGAATAGTTTCCTGCTTTAAGCAAAAGAC
CTACAATAGGTGGTGGAATTATGGGATGGGGTGGAGTATTGATATAAATATATAAATACAAATGTATATT
TTTCAGGATGTGGTTTAGGAACTGGGAATAACGTTTTCTGTTACTCCTGATGGTGCCATGAAAAGGTTAT
GTAATAAAATATTTTAAAATCAAAAAAAAAAAAAAAAAA

Human SNX1 protein sequence - var1 (public gi: 23111032) (SEQ ID NO: 281)

MASGGGGCSASERLPPPFPGLEPESEGAAGGSEPEAGDSDTEGEDIFTGAAVVSKHQSPKITTSLLPINN
GSKENGIHEEQDQEPQDLFAGDGMNAYVAYKVTTQTSLPLFRSKQFAVKRRFSDFLGLYEKLSEKHSQNG
FIVPPPPEKSLIGMTKVKVGKEDSSSAEFLEKRRAALERYLQRIVNHPTMLQDPDVREFLEKEELPRAVG
TQTLSGAGLLKMFNKATDAVSKMTIKMNESDIWFEEKLQEVECEEQRLRKLHAVVETLVNHRKELALNTA
QFAKSLAMLGSSEDNTALSRALSQLAEVEEKIEQLHQEQANNDFFLLAELLSDYIRLLAIVRAAFDQRMK
TWQRWQDAQATLQKKREAEARLLWANKPDKLQQAKDEILEWESRVTQYERDFERISTVVRKEVIRFEKEK
SKDFKNHVIKYLETLLYSQQQLAKYWEAFLPEAKAIS

Human SNX1 protein sequence - var2 (public gi: 23111036) (SEQ ID NO: 282)
MASGGGGCSASERLPPPFPGLEPESEGAAGGSEPEAGDSDTEGEDIFTGAAVVSKHQSPKITTSLLPINN
GSKENGIHEEQDQEPQDLFADATVELSLDSTQNNQKKVLAKTLISLPPQEATNSSKPQPTYEELEEEEQE
DQFDLTVGITDPEKIGDGMNAYVAYKVTTQTSLPLFRSKQFAVKRRFSDFLGLYEKLSEKHSQNGFIVPP
PPEKSLIGMTKVKVGKEDSSSAEFLEKRRAALERYLQRIVNHPTMLQDPDVREFLEKEELPRAVGTQTLS
GAGLLKMFNKATDAVSKMTIKMNESDIWFEEKLQEVECEEQRLRKLHAVVETLVNHRKELALNTAQFAKS
LAMLGSSEDNTALSRALSQLAEVEEKIEQLHQEQANNDFFLLAELLSDYIRLLAIVRWESRVTQYERDFE
RISTVVRKEVIRFEKEKSKDFKNHVIKYLETLLYSQQQLAKYWEAFLPEAKAIS

Human SNX1 protein sequence - var3 (public gi: 12653179) (SEQ ID NO: 283)
MASGGGGCSASERLPPPFPGLEPESEGAAGGSEPEAGDSDTEGEDIFTGAAVVSKHQSPKITTSLLPINN
GSKENGIHEEQDQEPQDLFADATVELSLDSTQNNQKKVLAKTLISLPPQEATNSSKPQPTYEELEEEEQE
DQFDLTVGITDPEKIGDGMNAYVAYKVTTQTSLPLFRSKQFAVKRRFSDFLGLYEKLSEKHSQNGFIVPP
PPEKSLIGMTKVKVGKEDSSSAEFLEKRRAALERYLQRIVNHPTMLQDPDVREFLEKEELPRAVGTQTLS
GAGLLKMFNKATDAVSKMTIKMNESDIWFEEKLQEVECEEQRLRKLHAVVETLVNHRKELALNTAQFAKS
LAMLGSSEDNTALSRALSQLAEVEEKIEQLHQEQANNDFFLLAELLSDYIRLLAIVRAAFDQRMKTWQRW
QDAQATLQKKREAEARLLWANKPDKLQQAKDEILEWESRVTQYERDFERISTVVRKEVIRFEKEKSKDFK
NHVIKYLETLLYSQQQLAKYWEAFLPEAKAIS

Human SNX1 protein sequence - var4 (public gi: 34535423) (SEQ ID NO: 284)
MASGGGGCSASERLPPPFPGLEPESEGAAGGSEPEAGDSDTEGEDIFTGAAVVSKHQSPKITTSLLPINN
GSKENGIHEEQDQEPQDLFADATVELSLDSTQNNQKKVLAKTLISLPPQEATNSSKPQPTYEELEEEEQE
DQFDLTVGITDPEKIGDGMNAYVAYKVTTQTSLPLFRSKQFAVKRRFSDFLGLYEKLSEKHSQNGFIVPP
PPEKSLIGMTKVKVGKEDSSSAEFLEKRRAALERYLQRIVNHPTMLQDPDVREFLEKEELPRAVGTQTLS
GAGLLKMFNKATDAVSKMTIKMNESDIWFEEKLQEVECEEQRLRKLHAVVETLVNHRKELALNTAQFAKS
LAMLGSSEDNTALSRALSQLAEVEEKIEQLHQEQANNDFFLLAELLSDYIRLLAIVRAAFDQRMKTWQRW
QDAQATLQKKREAEARLLWANKPDKLQQAKDEILEWESRVTQYERDFERISTVVRKEVIRFEKEKSKDFK
NHVIKYLETLLCSQQQAGEQLGIRSGILLTKKLPRYSKFFSTVHKFCAAASLWKWGFFLSAYLSYLF

Human SNX1 protein sequence - var5 (public gi: 3152942) (SEQ ID NO: 285)
MASGGGGCSASERLPPPFPGLEPESEGAAGGSEPEAGDSDTEGEDIFTGAAVVSKHQSPKITTSLLPINN
GSKENGIHEEQDQEPQDLFAGDGMNAYVAYKVTTQTSLPLFRSKQFAVKRRFSDFLGLYEKLSEKHSQNG
FIVPPSPEKSLIGMTKVKVGKEDSSSAEFLEKRRAALERYLQRIVNHPTMLQDPDVREFLEKEELPRAVG
TQTLSGAGLLKMFNKATDAVSKMTIKMNESDIWFEEKLQEVECEEQRLRKLHAVVETLVNHRKELALNTA
QFAKSLAMLGSSEDNTALSRALSQLAEVEEKIEQLHQEQANNDFFLLAELLSDYIRLLAIVRAAFDQRMK
TWQRWQDAQATLQKKREAEARLLWANKPDKLQQAKDEILEWESRVTQYERDFERISTVVRKEVIRFEKEK
SKDFKNHVIKYLETLLYSQQQLAKYWEAFLPEAKAIS

Human SNX1 protein sequence - var6 (public gi: 3152940) (SEQ ID NO: 286)
MASGGGGCSASERLPPPFPGLEPESEGAAGGSEPEAGDSDTEGEDIFTGAAVVSKHQSPKITTSLLPINN
GSKENGIHEEQDQEPQDLFADATVELSLDSTQNNQKKVLAKTLISLSPQEATNSSKPQPTYEELEEEEQE
DQFDLTVGITDPEKIGDGMNAYVAYKVTTQTSLPLFRSKQFAVKRRFSDFLGLYEKLSEKHSQNGFIVPP
SPEKSLIGMTKVKVGKEDSSSAEFLEKRRAALERYLQRIVNHPTMLQDPDVREFLEKEELPRAVGTQTLS
GAGLLKMFNKATDAVSKMTIKMNESDIWFEEKLQEVECEEQRLRKLHAVVETLVNHRKELALNTAQFAKS
LAMLGSSEDNTALSRALSQLAEVEEKIEQLHQEQANNDFFLLAELLSDYIRLLAIVRAAFDQRMKTWQRW
QDAQATLQKKREAEARLLWANKPDKLQQAKDEILEWESRVTQYERDFERISTVVRKEVIRFEKEKSKDFK
NHVIKYLETLLYSQQQLAKYWEAFLPEAKAIS

Unigene Name: SNX3 **Unigene ID:** Hs.12102

Human SNX3 mRNA sequence - var1 (public gi: 23111040) (SEQ ID NO: 162)

CTGTTTGCGACCCCGAGTCCCATGACACCGCTTCTCCTCACACCCCAGTCCGCAGTGCCCCTCCCCAGCC
TCGGCCGGGCCTCCCGGGAGCCGGGCGTGGCGTTCCAGCTAGTGAGCCGTTTCTCCCCTGGGCTCGGAGG
CGGAAGCTTGAGGGGCGCGGGGAGGAGCTTCGCGTGCGGGGTGAACGCCCGCTCTACGTGCTCGTTCTCT
TCGCGACCGCTGCGCGCGAGCCCCGTGTCCCCACGGCGGGCAGCAGCGGCGGCGGCGGCGGCTGAACGCG
GAGGGGGCGGAGGGAGCCCGCGGCGGCGGCAGCAGCTACAGCGAAATGGCGGAGACCGTGGCTGACACCC
GGCGGCTGATCACCAAGCCGCAGAACCTGAATGACGCCTACGGACCCCCCAGCAACTTCCTCGAGATCGA
TGTGAGCAACCCGCAAACGGTGGGGGTCGGCCGGGGCCGCTTCACCACTTACGAAATCAGGGTCAAGGTC
GTAGTTCCCCCGCTCCCTGGGAAAGCGTTTTTGCGTCAGCTTCCTTTTAGAGGAGATGATGGAATATTTG
ATGACAATTTTATTGAGGAAAGAAAACAAGGGCTGGAGCAGTTTATAAACAAGGTCGCTGGTCATCCTCT
GGCACAGAACGAACGTTGTCTTCACATGTTTTTACAAGATGAAATAATAGATAAAAGCTATACTCCATCT
AAAATAAGACATGCCTGAAATTTGGCAAGAAGGGGCAAAAACGTGACTATTAATGATTGATAAGCACCAG
TGAAGAAGTTCTAACTTTTAGCATGCTGCACAGAAACTGGTATAACATGCCTTCAGTATACTAACACTCA
TATGCTCAGTTTTGTTTTGTTTTGGCAGTTGACAAGAAGTTAATTTGCTTTAGTAAAAATCCCTCATTCC
AGCCTTTCTATATAAATAGCTCTTTCTTGCTGTTTTAATGTGGTGCACACTATAGCCTCACAAACCTGTT
ATTCCAGTGTAATCTGCAGTGTCGTAACTAAAGTTACTGGCTTGGTCTTATTTGCACAGTTTTTGCGTCT
TGTTTGCTTCTTGCATCTGATTAACTAGAATATTTCTCTTTCCCCCTTTTAATTTGTGATGTCACTTGAC
CCCATTTATGTGTAGGAGCACTACACCATTGGTTTCCAATACTGCACACATAAGATACATACTTGTGTGC
AGAAAGTATCTTCCTCCAGGCTTGTAATACCCTTCACATGGAAGATTAATGAGGGAAATCTTTATATTCT
GTATAAAAACAAAAGCAAATTTATATACTAAAATCATTTGTCTAAAAATTTAAGTTGTTTTCAAATAAAA
ATTAAAATGCATTTCTGATATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human SNX3 mRNA sequence - var2 (public gi: 34304375) (SEQ ID NO: 163)
GTCCGGCCGGAACCTGTTTGCGACCCCGAGTCCCATGACACCGCTTCTCCTCACACCCCAGTCCGCAGTG
CCCCTCCCCAGCCTCGGCCGGGCCTCCCGGGAGCCGGGCGTGGCGTTCCAGCTAGTGAGCCGTTTCTCCC
CTGGGCTCGGAGGCGGAAGCTTGAGGGGCGCGGGGAGGAGCTTCGCGTGCGGGGTGAACGCCCGCTCTAC
GTGCTCGTTCTCTTCGCGACCGCTGCGCGCGAGCCCCGTGTCCCCACGGCGGGCAGCAGCGGCGGCGGCG
GCGGCTGAACGCGGAGGGGGCGGAGGGAGCCCGCGGCGGCGGCAGCAGCTACAGCGAAATGGCGGAGACC
GTGGCTGACACCCGGCGGCTGATCACCAAGCCGCAGAACCTGAATGACGCCTACGGACCCCCCAGCAACT
TCCTCGAGATCGATGTGAGCAACCCGCAAACGGTGGGGGTCGGCCGGGGCCGCTTCACCACTTACGAAAT
CAGGGTCAAGACAAATCTTCCTATTTTCAAGCTGAAAGAATCTACTGTTAGAAGAAGATACAGTGACTTT
GAATGGCTGCGAAGTGAATTAGAAAGAGAGAGCAAGCCCTGCCTCAGAATGACATCAGAGGCAAGGAGTC
ATGGAAGGACGTGGTGTGCTCAGAATGATGAAAAGTTATTTTGTGACTAGAAAGTCGTAGTTCCCCCGCT
CCCTGGGAAAGCGTTTTTGCGTCAGCTTCCTTTTAGAGGAGATGATGGAATATTTGATGACAATTTTATT
GAGGAAAGAAAACAAGGGCTGGAGCAGTTTATAAACAAGGTCGCTGGTCATCCTCTGGCACAGAACGAAC
GTTGTCTTCACATGTTTTTACAAGATGAAATAATAGATAAAAGCTATACTCCATCTAAAATAAGACATGC
CTGAAATTTGGCAAGAAGGGGCAAAAACGTGACTATTAATGATTGATAAGCACCAGTGAAGAAGTTCTAA
CTTTTAGCATGCTGCACAGAAACTGGTATAACATGCCTTCAGTATACTAACACTCATATGCTCAGTTTTG
TTTTGTTTTGGCAGTTGACAAGAAGTTAATTTGCTTTAGTAAAAATCCCTCATTCCAGCCTTTCTATATA
AATAGCTCTTTCTTGCTGTTTTAATGTGGTGCACACTATAGCCTCACAAACCTGTTATTCCAGTGTAATC
TGCAGTGTCGTAACTAAAGTTACTGGCTTGGTCTTATTTGCACAGTTTTTGCGTCTTGTTTGCTTCTTGC
ATCTGATTAACTAGAATATTTCTCTTTCCCCCTTTTAATTTGTGATGTCACTTGACCCCATTTATGTGTA
GGAGCACTACACCATTGGTTTCCAATACTGCACACATAAGATACATACTTGTGTGCAGAAAGTATCTTCC
TCCAGGCTTGTAATACCCTTCACATGGAAGATTAATGAGGGAAATCTTTATATTCTGTATAAAAACAAAA
GCAAATTTATATACTAAAATCATTTGTCTAAAAATTTAAGTTGTTTTCAAATAAAAATTAAAATGCATTT
CTGATATGCAAAAAAAAAAAAAAAAAAAAAAA

Human SNX3 mRNA sequence - var3 (public gi: 34190889) (SEQ ID NO: 164)
TCGACCCACGCGTCCGCCCACGCGTCCGCTGTTTGCGACCCCGAGTCCCATGACACCGCTTCTCCTCACA
CCCCAGTCCGCAGTGCCCCTCCCCAGCCTCGGCCGGGCCTCCCGGGAGCCGGGCGTGGCGTTCCAGCTAG
TGAGCCGTTTCTCCCCTGGGCTCGGAGGCGGAAGCTTGAGGGGCGCGGGGAGGAGCTTCGCGTGCGGGGT
GAACGCCCGCTCTACGTGCTCGTTCTCTTCGCGACCGCTGCGCGCGAGCCCCGTGTCCCCACGGCGGGCA
GCAGCGGCGGCGGCGGCGGCTGAACGCGGAGGGGGCGGAGGGAGCCCGCGGCGGCGGCAGCAGCTACAGC
GAAATGGCGGAGACCGTGGCTGACACCCGGCGGCTGATCACCAAGCCGCAGAACCTGAATGACGCCTACG
GACCCCCCAGCAACTTCCTCGAGATCGATGTGAGCAACCCGCAAACGGTGGGGGTCGGCCGGGGCCGCTT
CACCACTTACGAAATCAGGGTCAAGACAAATCTTCCTATTTTCAAGCTGAAAGAATCTACTGTTAGAAGA
AGATACAGTGACTTTGAATGGCTGCGAAGTGAATTAGAAAGAGAGAGCAAGCCCTGCCTCAGAATGACAT
CAGAGGCAAGGAGTCATGGAAGGACGTGGTGTGCTCAGAATGATGAAAAGTTATTTTGTGACTAGAAAGT
CGTAGTTCCCCCGCTCCCTGGGAAAGCGTTTTTGCGTCAGCTTCCTTTTAGAGGAGATGATGGAATATTT
GATGACAATTTTATTGAGGAAAGAAAACAAGGGCTGGAGCAGTTTATAAACAAGGTCGCTGGTCATCCTC
TGGCACAGAACGAACGTTGTCTTCACATGTTTTTACAAGATGAAATAATAGATAAAAGCTATACTCCATC
TAAAATAAGACATGCCTGAAATTTGGCAAGAAGGGGCAAAAACGTGACTATTAATGATTGATAAGCACCA
GTGAAGAAGTTCTAACTTTTAGCATGCTGCACAGAAACTGGTATAACATGCCTTCAGTATACTAACACTC

ATATGCTCAGTTTTGTTTTGTTTTGGCAGTTGACAAGAAGTTAATTTGCTTTAGTAAAAATCCCTCATTC
CAGCCTTTCTATATAAATAGCTCTTTCTTGCTGTTTTAATGTGGTGCACACTATAGCCTCACAAACCTGT
TATTCCAGTGTAATCTGCAGTGTCGTAACTAAAGTTACTGGCTTGGTCTTATTTGCACAGTTTTTGCGTC
TTGTTTGCTTCTTGCATCTGATTAACTAGAATATTTCTCTTTCCCCCTTTTAATTTGTGATGTCACTTGA
CCCCATTTATGTGTAGGAGCACTACACCATTGGTTTCCAATACTGCACACATAAGATACATACTTGTGTG
CAGAAAGTATCTTCCTCCAGGCTTGTAATACCCTTCACATGGAAGATTAATGAGGGAAATCTTTATATTC
TGTATAAAAACAAAAGCAAATTTATATACTAAAATCATTTGTCTAAAAATTTAAGTTGTTTTCAAATAAA
AATTAAAATGCATTTCTGATATGCAAAAAAAAAAAAAAAAAAAAAA

Human SNX3 mRNA sequence - var4 (public gi: 15779011) (SEQ ID NO: 165)
GGGGCTTCGCGACCGCTGCGCGCGAGCCCCGTGTCCCCACGGCGGGCAGCAGCGGCGGCGGCGGCGGCTG
AACGCGGAGGGGGCGGAGGGAGCCCGCGGCGGCGGCAGCAGCTACAGCGAAATGGCGGAGACCGTGGCTG
ACACCCGGCGGCTGATCACCAAGCCGCAGAACCTGAATGACGCCTACGGACCCCCCAGCAACTTCCTCGA
GATCGATGTGAGCAACCCGCAAACGGTGGGGGTCGGCCGGGGCCGCTTCACCACTTACGAAATCAGGGTC
AAGACAAATCTTCCTATTTTCAAGCTGAAAGAATCTACTGTTAGAAGAAGATACAGTGACTTTGAATGGC
TGCGAAGTGAATTAGAAAGAGAGAGCAAGGTCGTAGTTCCCCCGCTCCCTGGAAAGCGTTTTTGCGTCA
GCTTCCTTTTAGAGGAGATGATGGAATATTTGATGACAATTTTATTGAGGAAAGAAAACAAGGGCTGGAG
CAGTTTATAAACAAGGTCGCTGGTCATCCTCTGGCACAGAACGAACGTTGTCTTCACATGTTTTTACAAG
ATGAAATAATAGATAAAAGCTATACTCCATCTAAAATAAGACATGCCTGAAATTTGGCAAGAAGGGGCAA
AAACGTGACTATTAATGATTGATAAGCACCAGTGAAGAAGTTCTAACTTTTAGCATGCTGCACAGAAACT
GGTATAACATGCCTTCAGTATACTAACACTCATATGCTCAGTTTTGTTTTGTTTTGGCAGTTGACAAGAA
GTTAATTTGCTTTAGTAAAAATCCCTCATTCCAGCCTTTCTATATAAATAGCTCTTTCTTGCTGTTTTAA
TGTGGTGCACACTATAGCCTCACAAACCTGTTATTCCAGTGTAATCTGCAGTGTCGTAACTAAAGTTACT
GGCTTGGTCTTATTTGCACAGTTTTTGCGTCTTGTTTGCTTCTTGCATCTGATTAACTAGAATATTTCTC
TTTCCCCCTTTTAATTTGTGATGTCACTTGACCCCATTTATGTGTAGGAGCACTACACCATTGGTTTCCA
ATACTGCACACATAAGATACATACTTGTGTGCAGAAAGTATCTTCCTCCAGGCTTGTAATACCCTTCACA
TGGAAGATTAATGAGGGAAATCTTTATATTCTGTATAAAAACAAAAGCAAATTTATATACTAAAATCATT
TGTCTAAAAATTTAAGTTGTTTTCAAATAAAAATTAAAATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

Human SNX3 mRNA sequence - var5 (public gi: 15929496) (SEQ ID NO: 166)
CGCGCGAGCCCCGTGTCCCCACGGCGGGCAGCAGCGGCGGCGGCGGCGGCTGAACGCGGAGGGGGCGGAG
GGAGCCCGCGGCGGCGGCAGCAGCTACAGCGAAATGGCGGAGACCGTGGCTGACACCCGGCGGCTGATCA
CCAAGCCGCAGAACCTGAATGACGCCTACGGACCCCCCAGCAACTTCCTCGAGATCGATGTGAGCAACCC
GCAAACGGTGGGGGTCGGCCGGGGCCGCTTCACCACTTACGAAATCAGGGTCAAGACAAATCTTCCTATT
TTCAAGCTGAAAGAATCTACTGTTAGAAGAAGATACAGTGACTTTGAATGGCTGCGAAGTGAATTAGAAA
GAGAGAGCAAGGTCGTAGTTCCCCCGCTCCCTGGAAAGCGTTTTTGCGTCAGCTTCCTTTTAGAGGAGA
TGATGGAATATTTGATGACAATTTTATTGAGGAAAGAAAACAAGGGCTGGAGCAGTTTATAAACAAGGTC
GCTGGTCATCCTCTGGCACAGAACGAACGTTGTCTTCACATGTTTTTACAAGATGAAATAATAGATAAAA
GCTATACTCCATCTAAAATAAGACATGCCTGAAATTTGGCAAGAAGGGGCAAAAACGTGACTATTAATGA
TTGATAAGCACCAGTGAAGAAGTTCTAACTTTTAGCATGCTGCACAGAAACTGGTATAACATGCCTTCAG
TATACTAACACTCATATGCTCAGTTTTGTTTTGTTTTGGCAGTTGACAAGAAGTTAATTTGCTTTAGTAA
AAATCCCTCATTCCAGCCTTTCTATATAAATAGCTCTTTCTTGCTGTTTTAATGTGGTGCACACTATAGC
CTCACAAACCTGTTATTCCAGTGTAATCTGCAGTGTCGTAACTAAAGTTACTGGCTTGGTCTTATTTGCA
CAGTTTTTGCGTCTTGTTTGCTTCTTGCATCTGATTAACTAGAATATTTCTCTTTCCCCCTTTTAATTTG
TGATGTCACTTGACCCCATTTATGTGTAGGAGCACTACACCATTGGTTTCCAATACTGCACACATAAGAT
ACATACTTGTGTGCAGAAAGTATCTTCCTCCAGGCTTGTAATACCCTTCACATGGAAGATTAATGAGGGA
AATCTTTATATTCTGTATAAAAACAAAAGCAAATTTATATACTAAAATCATTTGTCTAAAAATTTAAGTT
GTTTTCAAATAAAAATTAAAATGCATTTCTGATATGCAAAAAAAAAAAAAAA

Human SNX3 mRNA sequence - var6 (public gi: 14250078) (SEQ ID NO: 167)
AGCCCCGTGTCCCCACGGCGGGCAGCAGCGGCGGCGGCGGCGGCTGAACGCGGAGGGGGCGGAGGGAGCC
CGCGGCGGCGGCAGCAGCTACAGCGAAATGGCGGAGACCGTGGCTGACACCCGGCGGCTGATCACCAAGC
CGCAGAACCTGAATGACGCCTACGGACCCCCCAGCAACTTCCTCGAGATCGATGTGAGCAACCCGCAAAC
GGTGGGGGTCGGCCGGGGCCGCTTCACCACTTACGAAATCAGGGTCAAGACAAATCTTCCTATTTTCAAG
CTGAAAGAATCTACTGTTAGAAGAAGATACAGTGACTTTGAATGGCTGCGAAGTGAATTAGAAAGAGAGA
GCAAGGTCGTAGTTCCCCCGCTCCCTGGAAAGCGTTTTTGCGTCAGCTTCCTTTTAGAGGAGATGATGG
AATATTTGATGACAATTTTATTGAGGAAAGAAAACAAGGGCTGGAGCAGTTTATAAACAAGGTCGCTGGT
CATCCTCTGGCACAGAACGAACGTTGTCTTCACATGTTTTTACAAGATGAAATAATAGATAAAAGCTATA
CTCCATCTAAAATAAGACATGCCTGAAATTTGGCAAGAAGGGGCAAAAACGTGACTATTAATGATTGATA
AGCACCAGTGAAGAAGTTCTAACTTTTAGCATGCTGCACAGAAACTGGTATAACATGCCTTCAGTATACT
AACACTCATATGCTCAGTTTTGTTTTGTTTTGGCAGTTGACAAGAAGTTAATTTGCTTTAGTAAAAATCC

CTCATTCCAGCCTTTCTATATAAATAGCTCTTTCTTGCTGTTTTAATGTGGTGCACACTATAGCCTCACA
AACCTGTTATTCCAGTGTAATCTGCAGTGTCGTAACTAAAGTTACTGGCTTGGTCTTATTTGCACAGTTT
TTGCGTCTTGTTTGCTTCTTGCATCTGATTAACTAGAATATTTCTCTTTCCCCCTTTTAATTTGTGATGT
CACTTGACCCCATTTATGTGTAGGAGCACTACACCATTGGTTTCCAATACTGCACACATAAGATACATAC
TTGTGTGCAGAAAGTATCTTCCTCCAGGCTTGTAATACCCTTCACATGGAAGATTAATGAGGGAAATCTT
TATATTCTGTATAAAAACAAAAGCAAATTTATATACTAAAATCATTTGTCTAAAAATTTAAGTTGTTTTC
AAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human SNX3 mRNA sequence - var7 (public gi: 12957159) (SEQ ID NO: 168)
GGGCGAGGAGGGAGCCCGCGGCGGCGGCAGCAGCTACAGCGAAATGGCGGAGACCGTGGCTGACACCCGG
CGGCTGATCACCAAGCCGCAGAACCTGAATGACGCCTACGGACCCCCCAGCAACTTCCTCGAGATCGATG
TGAGCAACCCGCAAACGGTGGGGGTCGGCCGGGGCCGCTTCACCACTTACGAAATCAGGGTCAAGGTCGT
AGTTCCCCCGCTCCCTGGGAAAGCGTTTTTGCGTCAGCTTCCTTTTAGAGGAGATGATGGAATATTTGAT
GACAATTTTATTGAGGAAAGAAAACAAGGGCTGGAGCAGTTTATAAACAAGGTCGCTGGTCATCCTCTGG
CACAGAACGAACGTTGTCTTCACATGTTTTTACAAGATGAAATAATAGATAAAAGCTATACTCCATCTAA
AATAAGACATGCCTGAAATTTGGCAAGAAGGGGCAAAAACGTGACTATTAATGATTGATAAGCACCAGTG
AAGAAGTTCTAACTTTTAGCATGCTGCACAGAAACTGGTATAACATGCCTTCAGTATACTAACACTCATA
TGCTCAGTTTTGTTTTGTTTTGGCAGTTGACAAGAAGTTAATTTGCTTTAGTAAAAATCCCTCATTCCAG
CCTTTCTATATAAATAGCTCTTTCTTGCTGTTTTAATGTGGTGCACACTATAGCCTCACAAACCTGTTAT
TCCAGTGTAATCTGCAGTGTCGTAACTAAAGTTACTGGCTTGGTCTTATTTGCACAGTTTTTGCGTCTTG
TTTGCTTCTTGCATCTGATTAACTAGAATATTTCTCTTTCCCCCTTTTAATTTGTGATGTCACTTGACCC
CATTTATGTGTAGGAGCACTACACCATTGGTTTCCAATACTGCACACATAAGATACATACTTGTGTGCAG
AAAGTATCTTCCTCCAGGCTTGTAATACCCTTCACATGGAAGATTAATGAGGGAAATCTTTATATTCTGT
ATAAAAACAAAAGCAAATTTATATACTAAAATCATTTGTCTAAAAATTTAAGTTGTTTTCAAATAAAAAT
TAAAATGCATTTCTGATATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human SNX3 mRNA sequence - var8 (public gi: 34304374) (SEQ ID NO: 169)
GTCCGGCCGGAACCTGTTTGCGACCCCGAGTCCCATGACACCGCTTCTCCTCACACCCCAGTCCGCAGTG
CCCCTCCCCAGCCTCGGCCGGGCCTCCCGGGAGCCGGGCGTGGCGTTCCAGCTAGTGAGCCGTTTCTCCC
CTGGGCTCGGAGGCGGAAGCTTGAGGGGCGCGGGGAGGAGCTTCGCGTGCGGGGTGAACGCCCCGCTCTAC
GTGCTCGTTCTCTTCGCGACCGCTGCGCGCGAGCCCCGTGTCCCCACGGCGGGCAGCAGCGGCGGCGGCG
GCGGCTGAACGCGGAGGGGGCGGAGGGAGCCCGCGGCGGCGGCAGCAGCTACAGCGAAATGGCGGAGACC
GTGGCTGACACCCGGCGGCTGATCACCAAGCCGCAGAACCTGAATGACGCCTACGGACCCCCCAGCAACT
TCCTCGAGATCGATGTGAGCAACCCGCAAACGGTGGGGGTCGGCCGGGGCCGCTTCACCACTTACGAAAT
CAGGGTCAAGACAAATCTTCCTATTTTCAAGCTGAAAGAATCTACTGTTAGAAGAAGATACAGTGACTTT
GAATGGCTGCGAAGTGAATTAGAAAGAGAGAGCAAGGTCGTAGTTCCCCCGCTCCCTGGGAAAGCGTTTT
TGCGTCAGCTTCCTTTTAGAGGAGATGATGGAATATTTGATGACAATTTTATTGAGGAAAGAAAACAAGG
GCTGGAGCAGTTTATAAACAAGGTCGCTGGTCATCCTCTGGCACAGAACGAACGTTGTCTTCACATGTTT
TTACAAGATGAAATAATAGATAAAAGCTATACTCCATCTAAAATAAGACATGCCTGAAATTTGGCAAGAA
GGGGCAAAAACGTGACTATTAATGATTGATAAGCACCAGTGAAGAAGTTCTAACTTTTAGCATGCTGCAC
AGAAACTGGTATAACATGCCTTCAGTATACTAACACTCATATGCTCAGTTTTGTTTTGTTTTGGCAGTTG
ACAAGAAGTTAATTTGCTTTAGTAAAAATCCCTCATTCCAGCCTTTCTATATAAATAGCTCTTTCTTGCT
GTTTTAATGTGGTGCACACTATAGCCTCACAAACCTGTTATTCCAGTGTAATCTGCAGTGTCGTAACTAA
AGTTACTGGCTTGGTCTTATTTGCACAGTTTTTGCGTCTTGTTTGCTTCTTGCATCTGATTAACTAGAAT
ATTTCTCTTTCCCCCTTTTAATTTGTGATGTCACTTGACCCCATTTATGTGTAGGAGCACTACACCATTG
GTTTCCAATACTGCACACATAAGATACATACTTGTGTGCAGAAAGTATCTTCCTCCAGGCTTGTAATACC
CTTCACATGGAAGATTAATGAGGGAAATCTTTATATTCTGTATAAAAACAAAAGCAAATTTATATACTAA
AATCATTTGTCTAAAAATTTAAGTTGTTTTCAAATAAAAATTAAAATGCATTTCTGATATGCAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

Human SNX3 mRNA sequence - var9 (public gi: 30583066) (SEQ ID NO: 170)
ATGGCGGAGACCGTGGCTGACACCCGGCGGCTGATCACCAAGCCGCAGAACCTGAATGACGCCTACGGAC
CCCCCAGCAACTTCCTCGAGATCGATGTGAGCAACCCGCAAACGGTGGGGGTCGGCCGGGGCCGCTTCAC
CACTTACGAAATCAGGGTCAAGACAAATCTTCCTATTTTCAAGCTGAAAGAATCTACTGTTAGAAGAAGA
TACAGTGACTTTGAATGGCTGCGAAGTGAATTAGAAAGAGAGAGCAAGGTCGTAGTTCCCCCGCTCCCTG
GGAAAGCGTTTTTGCGTCAGCTTCCTTTTAGAGGAGATGATGGAATATTTGATGACAATTTTATTGAGGA
AAGAAAACAAGGGCTGGAGCAGTTTATAAACAAGGTCGCTGGTCATCCTCTGGCACAGAACGAACGTTGT
CTTCACATGTTTTTACAAGATGAAATAATAGATAAAAGCTATACTCCATCTAAAATAAGACATGCCTAG

Human SNX3 mRNA sequence - var10 (public gi: 3127052) (SEQ ID NO: 171)
GGGCGAGGAGGGAGCCCGCGGCGGCGGCAGCAGCTACAGCGAAATGGCGGAGACCGTGGCTGACACCCGG

CGGCTGATCACCAAGCCGCAGAACCTGAATGACGCCTACGGACCCCCCAGCAACTTCCTCGAGATCGATG
TGAGCAACCCGCAAACGGTGGGGGTCGGCCGGGGCCGCTTCACCACTTACGAAATCAGGGTCAAGACAAA
TCTTCCTATTTTCAAGCTGAAAGAATCTACTGTTAGAAGAAGATACAGTGACTTTGAATGGCTGCGAAGT
GAATTAGAAAGAGAGAGCAAGGTCGTAGTTCCCCCGCTCCCTGGGAAAGCGTTTTTGCGTCAGCTTCCTT
TTAGAGGAGATGATGGAATATTTGATGACAATTTTATTGAGGAAAGAAAACAAGGGCTGGAGCAGTTTAT
AAACAAGGTCGCTGGTCATCCTCTGGCACAGAACGAACGTTGTCTTCACATGTTTTTACAAGATGAAATA
ATAGATAAAAGCTATACTCCATCTAAAATAAGACATGCCTGAAATTTGGCAAGAAGGGGCAAAAACGTGA
CTATTAATGATTGATAAGCACCAGTGAAGAAGTTCTAACTTTTAGCATGCTGCACAGAAACTGGTATAAC
ATGCCTTCAGTATACTAACACTCATATGCTCAGTTTTGTTTTGTTTTGGCAGTTGACAAGAAGTTAATTT
GCTTTAGTAAAAATCCCTCATTCCAGCCTTTCTATATAAATAGCTCTTTCTTGCTGTTTTAATGTGGTGC
ACACTATAGCCTCACAAACCTGTTATTCCAGTGTAATCTGCAGTGTCGTAACTAAAGTTACTGGCTTGGT
CTTATTTGCACAGTTTTTGCGTCTTGTTTGCTTCTTGCATCTGATTAACTAGAATATTTCTCTTTCCCCC
TTTTAATTTGTGATGTCACTTGACCCCATTTATGTGTAGGAGCACTACACCATTGGTTTCCAATACTGCA
CACATAAGATACATACTTGTGTGCAGAAAGTATCTTCCTCCAGGCTTGTAATACCCTTCACATGGAAGAT
TAATGAGGGAAATCTTTATATTCTGTATAAAAACAAAAGCAAATTTATATACTAAAATCATTTGTCTAAA
AATTTAAGTTGTTTTCAAATAAAAATTAAAATGCATTTCTGATATGCAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAA

Human SNX3 mRNA sequence - var11 (public gi: 3126978) (SEQ ID NO: 172)
GCGGCACAGCTACAGCGAAATGGCGGAGACCGTGGCTGACACCCGGCGGCTGATCACCAAGCCGCAGAAC
CTGAATGACGCCTACGGACCCCCCAGCAACTTCCTCGAGATCGATGTGAGCAACCCGCAAACGGTGGGGG
TCGGCCGGGGCCGCTTCACCACTTACGAAATCAGGGTCAAGACAAATCTTCCTATTTTCAAGCTGAAAGA
ATCTACTGTTAGAAGAAGATACAGTGACTTTGAATGGCTGCGAAGTGAATTAGAAAGAGAGAGCAAGGTC
GTAGTTCCCCCGCTCCCTGGGAAAGCGTTTTTGCGTCACTTCCCTTTTAGAGGAGATGATGGAATATTTG
ATGACAATTTTATTGAGGAAAGAAAACAAGGGCTGGAGCAGTTTATAAACAAGGTCGCTGGTCATCCTCT
GGCACAGAACGAACGTTGTCTTCACATGTTTTTACAAGATGAAATAATAGATAAAAGCTATACTCCATCT
AAAATAAGACATGCCTGAAATTTGGCAAGAAGGGGCAAAAACCGTGACTATTAATGATTGATAAGCACCA
GTGAAGAAGTTCTAACTTTTAGCATGCTGCACAGAAACTGGTATAACATGCCTTCAGTATACTAACACTC
CATATGCTCAGTTTTGTTTTGTTTTGGCAGTTGACAAGAAGTTAATTTGCTTTAGTAAAAATCCCTCATT
CCAGCCTTTCTATATAAATAGCTCCTTCCTGCTGTTTTAATGTGGGTGCACACTATAGCCTCACAACCTG
GTTAATCCAGTGTAATCTGCAGTGTCGTAACTAAAGTACTGGCTTGGTCCTAATTG

Human SNX3 protein sequence - var1 (public gi: 23111041) (SEQ ID NO: 287)
MAETVADTRRLITKPQNLNDAYGPPSNFLEIDVSNPQTVGVGRGRFTTYEIRVKVVVPPLPGKAFLRQLP
FRGDDGIFDDNFIEERKQGLEQFINKVAGHPLAQNERCLHMFLQDEIIDKSYTPSKIRHA

Human SNX3 protein sequence - var2 (public gi: 23111043) (SEQ ID NO: 288)
MAETVADTRRLITKPQNLNDAYGPPSNFLEIDVSNPQTVGVGRGRFTTYEIRVKTNLPIFKLKESTVRRR
YSDFEWLRSELERESKPCLRMTSEARSHGRTWCAQNDEKLFCD

Human SNX3 protein sequence - var3 (public gi: 15779012) (SEQ ID NO: 289)
MAETVADTRRLITKPQNLNDAYGPPSNFLEIDVSNPQTVGVGRGRFTTYEIRVKTNLPIFKLKESTVRRR
YSDFEWLRSELERESKVVVPPLPGKAFLRQLPFRGDDGIFDDNFIEERKQGLEQFINKVAGHPLAQNERC
LHMFLQDEIIDKSYTPSKIRHA

Human SNX3 protein sequence - var4 (public gi: 3126979) (SEQ ID NO: 290)
MAETVADTRRLITKPQNLNDAYGPPSNFLEIDVSNPQTVGVGRGRFTTYEIRVKTNLPIFKLKESTVRRR
YSDFEWLRSELERESKVVVPPLPGKAFLRHFPFRGDDGIFDDNFIEERKQGLEQFINKVAGHPLAQNERC
LHMFLQDEIIDKSYTPSKIRHA

Human SNX3 pray sequence - var1 (SEQ ID NO: 173)
GCCGCCATGGNAGTACCCATACGACGTACCAGATTACGCTCATATGGCCATGGAGGCCAGTGAATTCCAC
CCAAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCGGCGGCGGCGGCGGCTGAACGCGGAGGGGGCGG
AGGGAGCCCGCGGCGGCGGCAGCAGCTACAGCGAAATGGCGGAGACCGTGGCTGACACCCGGCGGCTGAT
CACCAAGCCGCAGAACCTGAATGACGCCTACGGACCCCCCAGCAACTTCCTCGAGATCGATGTGAGCAAC
CCGCAAACGGTGGGGGTCGGCCGGGGCCGCTTCACCACTTACGAAATCAGGGTCAAGACAAATCTTCCTA
TTTTCAAGCTGAAAGAATCTACTGTTAGAAGAAGATACAGTGACTTTGAATGGCTGCGAAGTGAATCAGA
AAGAGAGAGCAAGGTCGTAGTTCCCCNNGCTCCCTGGGAAAGCGTTTTTGCGTCAGCTTCCTTTTAGAGG
AGATGATGGAATATTTGATGACAATTTTATTGAGGAAAGAAAACAAGGGCTGGAGCAGTTTATAAACAAG
GTCGCTGGTCATCCTCTGGCACAAAACGAACGTTGTCTTCACATGTTTTTACANGATGAAATANTNGATA
AAAGCTNTACTCCATCTAAAATAAAACATGCCTGAANTTTGGCANAANGGGCNAAAACGTGACTATTATG

ATTGANAGCCCCNNNNAAAANTTCTANNTTTNNCNTGCTNACAAAACTGNNTAANTGCCTNANNTACTAA
CCTNNNTNCCNANTTTNNTTTGNNTGGNNNTNAAAAAATNAT

Human SNX3 pray sequence - var2 (SEQ ID NO: 174)
CCGCCATGGTAGTACCCATACGACGTACCAGTATTACGCTCATATGGCCATGGCAGGCCAGTGAATTCCA
CCCAAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCCGGGGGCAGGAGGGAGCCCAGCGGCGGCGGCA
GCAGCTACAGCGAAATGGCGGAGACCGTGGCTGACACCCGGCGGCTGATCACCAAGCCGCAGAACCTGAA
TGACGCCTACGGACCCCCCAGCAACTTCCTCGAGATCGATGTGAGCAACCCGCAAACGGTGGGGGTCGGC
CGGGGCCGCTTCACCACTTACGAAATCAGGGTCAAGACAAATCTTCCTATTTTCAAGCTGAAAGAATCTA
CTGTTAGAAGAAGATACAGTGACTTTGAATGGCTGCGAAGTGAATTAGAAAGAGAGAGCAAGGTCGTAGT
TCCCCCGCTCCCTGGGAAAGCGTTTTTGCGTCAGCTTNCTTTAGAGGGGATGATGGAATATTTGATGAC
AATTTTATTGAGGAAAGAAAACAAGGGCTGGANCANTTTATNAACAAGTNAGTGCTTNCTATTCCTNAAA
GTGTANGACTNCTTTAAGTGACTACTTTTNTTTANATGTNAANNNAACTGNACTGTNNCNTTTNTTTNAN
CNTTTCCTANNTTTNATTTNTTTAA

Unigene Name: SRA1 **Unigene ID:** Hs.32587 **Clone ID:** 3GD_19

Human SRA1 mRNA sequence - var1 (public gi: 10436964) (SEQ ID NO: 175)
ACGTGAAGCCGGGTGAGCGCAGCCGGCGGGCTAGGGCACTAGGTCGTCGCCCCGGCCTAGGCTGGGGGCG
GTTGCGGCGCTTAGTATGGACCCTCTGTCTCCCCCAGCCCCAGTATAAGCTAACAGTGGAGTTCCGGGCT
CGCTTCACACATCCCTCGCCTCCGCAGGCAACAAGGAACGCGGCTGGAACGACCCGCCGCAGTTCTCATA
CGGGCTGCAGACCCAGGCCGGCGGACCCAGGCGCTCGCTGCTTACCAAGAGGGTAGCCGCACCCCAGGAT
GGATCCCCCAGAGTCCCCGCATCAGAGACTTCTCCTGGGCCTCCCCCAATGGGGCCTCCACCTCCTTCAA
GTAAGGTTCCCAGGTCCCCACCTGTGGGGAGTGGTCCTGCCTCTGGCGTGGAGCCCACAAGTTTCCCAGT
CGAGTCTGAGGCTCGACTGATGGAGGATGTGCTGAGACCTTTGGAACAGGCATTGGAAGACTGCCGTGGC
CACACAAGGAAGCAGGTATGTGATGACATCAGCCGACGCCTGGCACTGCTGCAGGAACAGTGGGCTGGAG
GAAAGTTGTCAATACCTGTAAAGAAGAGAATGGCTCTACTGGTGCAAGAGCTTTCAAGCCACCGGTGGGA
CGCAGCAGATGACATCCACCGCTCCCTCATGGTTGACCATGTGACTGAGGTCAGTCAGTGGATGGTAGGA
GTTAAAAGATTAATTGCAGAAAAGAGGAGTCTGTTTTCAGAGGAGGCAGCCAATGAAGAGAAATCTGCAG
CCACAGCTGAGAAGAACCATACCATACCAGGCTTCCAGCAGGCTTCATAATCCTCGGTTCCCCAGACTCA
CCGGACACCATCTCCTATGCCTTGGAGACCTTCTGTCACTTGGCTCCCTTCTTACCACCACCAAGACTGT
CCCACTGGGCCTGACCCACCTATGAGGGAAGAAGTCCCACCTGGGCCAGAGGGAGTTCATGTGTTACTCA
TAACATGCATTTCAATAAAAACATCTCTGCGGTGGGCCTTGGGTAGGAGAGATGAACCCTTCCGGTGCCA
AGCTAGTCCCCTCTGGTGTCCTCGACTGCCCTGCTCCCTGTGTATCTGCAAACCTCTGTTCTCCCTTCTC
CATTCATCAGGAAGGGATCTGCTGGGTAAAGTCAGACTACTGCCTACCACTTTTTCCCAAAGTAGACTGA
AAGCACATCCTGTGCTGGGCGGAGCAGCTGTGTTTGGATGGTTTCATTTCAGCATGAGAACAGACTCAAA
TAGAACGGGGAGACTTTTCCCTCAACAAAAGGAAAGACAGTCCTATTTGCACTGTATCACCCTTGAGATA
CTACTGTTACAGAGATTAGAACCACATTGAGTGGGGTTTTCTGTGTAAATCGAAGGAGAAAAAGACCAGA
TTACTGAGATTGGGGATTGTAACTCTGACTTGCCAAACAAACTGCTGCCTCAAAAAAAAAAAAAAAAA

Human SRA1 mRNA sequence - var2 (public gi: 9930611) (SEQ ID NO: 176)
TCCTTTGGTGCCTTGTGACCAGGGCCCTGATGGTTCATTAGATGGAGCCTTCGAGTCTTAGGGAGTTGCC
GCAGGGTCCCCACAGCGGCTCCCGACGGTTGTGAACCAGCATCCATCCTCCACGGATTCCGGCAACCCGC
CTGGCCCTGGACGTGTCTCAACTGGCCCGCGTGAGGGGCCGCCCCGGAAATGACGCGCTGCCCCGCTGGC
CAAGCGGAAGTGGAGATGGCGGAGCTGTACGTGAAGCCGGGCAACAAGGAACGCGGCTGGAACGACCCGC
CGCAGTTCTCATACGGGCTGCAGACCCAGGCCGGCGGACCCAGGCGCTCGCTGCTTATCAAGAGGGTCGC
CGCACCCCAGGATGGATCCCCCAGAGTCCCCGCATCAGAGACTTCTCCTGGGCCTCCCCCAATGGGGCCT
CCACCTCCTTCAAGTAAGGCTCCCAGGTCCCCACCTGTGGGGAGTGGTCCTGCCTCTGGCGTGGAGCCCA
CAAGTTTCCCAGTCGAGTCTGAGGCTGTGATGGAGGATGTGCTGAGACCTTTGGAACAGGCATTGGAAGA
CTGCCGTGGCCACACAAGGAAGCAGGTATGTGATGACATCAGCCGACGCCTGGCACTGCTGCAGGAACAG
TGGGCTGGAGGAAAGTTGTCAATACCTGTAAAGAAGAGAATGGCTCTACTGGTGCAAGAGCTTTCAAGCC
ACCGGTGGGACGCAGCAGATGACATCCACCGCTCCCTCATGGTTGACCATGTGACTGAGGTCAGTCAGTG
GATGGTAGGAGTTAAAAGATTAATTGCAGAAAAGAGGAGTCTGTTTTCAGAGGAGGCAGCCAATGAAGAG
AAATCTGCAGCCACAGCTGAGAAGAACCATACCATACCAGGCTTCCAGCAGGCTTCATAATCCTCGGTTC
CCCAGACT

Human SRA1 mRNA sequence - var3 (public gi: 9930613) (SEQ ID NO: 177)
TCCTTTGGTGCCTTGTGACCAGGGCCCTGATGGTTCATTAGATGGAGCCTTCGAGTCTTAGGGAGTTGCC
GCAGGGTCCCCACAGCGGCTCCCGACGGTTGTGAACCAGCATCCATCCTCCACGGATTCCGGCAACCCGC
CTGGCCCTGGACGTGTCTCAACTGGCCCGCGTGAGGGGCCGCCCCGGAAATGACGCGCTGCCCCGCTGGC
CAAGCGGAAGTGGAGATGGCGGAGCTGTACGTGAAGCCGGGCAACAAGGAACGCGGCTGGAACGACCCGC

CGCAGTTCTCATACGGGCTGCAGACCCAGGCCGGCGGACCCAGGCGCTCGCTGCTTACCAAGAGGGTAGC
CGCACCCCAGGATGGATCCCCCAGAGTCCCCGCATCAGAGACTTCTCCTGGGCCTCCCCCAATGGGGCCT
CCACCTCCTTCAAGTAAGGCTCCCAGGTCCCCACCTGTGGGGAGTGGTCCTGCCTCTGGCGTGGAGCCCA
CAAGTTTCCCAGTCGAGTCTGAGGCTCGACTGATGGAGGATGTGCTGAGACCTTTGGAACAGGCATTGGA
AGACTGCCGTGGCCACACAAGGAAGCAGGTATGTGATGACATCAGCCGACGCCTGGCACTGCTGCAGGAA
CAGTGGGCTGGAGGAAAGTTGTCAATACCTGTAAAGAAGAGAATGGCTCTACTGGTGCAAGAGCTTTCAA
GCCACCGGTGGGACGCAGCAGATGACATCCACCGCTCCCTCATGGTTGACCATGTGACTGAGGTCAGTCA
GTGGATGGTAGGAGTTAAAAGATTAATTGCAGAAAAGAGGAGTCTGTTTTCAGAGGAGGCAGCCAATGAA
GAGAAATCTGCAGCCACAGCTGAGAAGAACCATACCATACCAGGCTTCCAGCAGGCTTCATAATCCTCGG
TTCCCCAGACT

Human SRA1 mRNA sequence - var4 (public gi: 4588026) (SEQ ID NO: 178)
CGCTTGGCGGAGCTGTACGTGAAGCCGGGCAACAAGGAACGCGGCTGGAACGACCCGCCGCAGTTCTCAT
ACGGGCTGCAGACCCAGGCCGGCGGACCCAGGCGCTCGCTGCTTACCAAGAGGGTAGCCGCACCCCAGGA
TGGATCCCCCAGAGTCCCCGCATCAGAGACTTCTCCTGGGCCTCCCCCAATGGGGCCTCCACCTCCTTCA
AGTAAGGCTCCCAGGTCCCCACCTGTGGGGAGTGGTCCTGCCTCTGGCGTGGAGCCCACAAGTTTCCCAG
TCGAGTCTGAGGCTGTGATGGAGGATGTGCTGAGACCTTTGGAACAGGCATTGGAAGACTGCCGTGGCCA
CACAAGGAAGCAGGTATGTGATGACATCAGCCGACGCCTGGCACTGCTGCAGGAACAGTGGGCTGGAGGA
AAGTTGTCAATACCTGTAAAGAAGAGAATGGCTCTACTGGTGCAAGAGCTTTCAAGCCACCGGTGGGACG
CAGCAGATGACATCCACCGCTCCCTCATGGTTGACCATGTGACTGAGGTCAGTCAGTGGATGGTAGGAGT
TAAAAGATTAATTGCAGAAAAGAGGAGTCTGTTTTCAGAGGAGGCAGCCAATGAAGAGAAATCTGCAGCC
ACAGCTGAGAAGAACCATACCATACCAGGCTTCCAGCAGGCTTCATAATCCTCGGTTCCCCAGACTCACC
GGACACCATCTCCTATGCCTTGGAGACCTTCTGTCACTTGGCTCCCTTCTTACCACCACCAAGACTGTCC
CACTGGGCCTGACCCACCTATGAGGGAAGAAGTCCCACCTGGGCCAGAGGGAGTTCATGTGTTACTCATA
ACATGCATTTCAATAAAAACATCTCTGCGGTGGTG

Human SRA1 mRNA sequence - var5 (public gi: 25123254) (SEQ ID NO: 179)
GGCGGAGCTGTACGTGAAGCCGGGCAACAAGGAACGCGGCTGGAACCCCGCCGCAGTTCTCATACGGGCT
GCAGACCCAGGCCGGCGGACCCAGGCGCTCGCTGCTTACCAAGAGGGTCGCCGCACCCCAGGATGGATCC
CCCAGAGTCCCCGCATCAGAGACTTCTCCTGGGCCTCCCCCAATGGGGCCTCCACCTCCTTCAAGTAAGG
CTCCCAGGTCCCCACCTGTGGGGAGTGGTCCTGCCTCTGGCGTGGAGCCCACAAGTTTCCCAGTCGAGTC
TGAGGCTGTGATGGAGGATGTGCTGAGACCTTTGGAACAGGCATTGGAAGACTGCCGTGGCCACACAAGG
AAGCAGGTATGTGATGACATCAGCCGACGCCTGGCACTGCTGCAGGAACAGTGGGCTGGAGGAAAGTTGT
CAATACCTGTAAAGAAGAGAATGGCTCTACTGGTGCAAGAGCTTTCAAGCCACCGGTGGGACGCAGCAGA
TGACATCCACCGCTCCCTCATGGTTGACCATGTGACTGAGGTCAGTCAGTGGATGGTAGGAGTTAAAAGA
TTAATTGCAGAAAAGAGGAGTCTGTTTTCAGAGGAGGCAGCCAATGAAGAGAAATCTGCAGCCACAGCTG
AGAAGAACCATACCATACCAGGCTTCCAGCAGGCTTCATAATCCTCGGTTCCCCAGACTCACCGGACACC
ATCTCCTATGCCTTGGAGACCTTCTGTCACTTGGCTCCCTTCTTACCACCACCAAGACTGTCCCACTGGG
CCTGACCCACCTATGAGGGAAGAAGTCCCACCTGGGCCAGAGGGAGTTCATGTGTTACTCATAACATGCA
TTTCAATAAAAACATCTCTGCGGTGAAAAAAAAAAAAAAA

Human SRA1 mRNA sequence - var6 (public gi: 18027813) (SEQ ID NO: 180)
GCAGGCACTAAGCTGGGCACTGGGAATGTAATAAAATAGTCAAGGTCCCACCTTCTAAGACTGTCCGACA
GGGAAACGAACAAGAGTCAAATAAGGCAGAAGATGTGATGTAATACACCTACGAAATCTCAGAGGGTTGT
AGGGTCGTGGGAGCTCAAGTGAGACACTTAACCTGGCCTGAGACATTCCAGAAGGCCTCCTGAAGAACTG
ACATCTGAACTGAGAACTGAAGGAAGATGAGTACTAGTGAGGCTACCGGACGTGAATGTGGAGATTGTGC
AGGGCAATGCAAGAGGAGGCTGTAGAAGTCAACCTGGCTAGATCACAGCGGGGTGTATGTGGGGCAGGAG
CTTCTTTGTTTGAATTTGCTCCTGAGAGGATGAGGCCTCCTAGAGCACTGGCTCCTGGACAGCAACCTCC
TTTGGTGCCTTGTGACCAGGGCCCTGATGGTTCATTAGATGGAGCCTTCGAGTCTTAGGGAGTTGCCGCA
GGGTCCCCACAGCGGCTCCCGACGGTTGTGAACCAGCATCCATTCTCCACGGATTCCGGCAACCCGCCTG
GCCCTGGACGTGTCTCAACTGGCCCGCGTGAGGGGCCGCCCCGGAAATGACGCGCTGCCCCGCTGGCCAA
GCGGAAGTGGAGATGGCGGAGCTGTACGTGAAGCCGGGCAACAAGGAACGCGGCTGGAACGACCCGCCGC
AGTTCTCATACGGGCTGCAGACCCAGGCCGGCGGACCCAGGCGCTCGCTGCTTACCAAGAGGGTAGCCGC
ACCCCAGGATGGATCCCCCAGAGTCCCCGCATCAGAGACTTCTCCTGGGCCTCCCCCAATGGGGCCTCCA
CCTCCTTCAAGTAAGGCTCCCAGGTCCCCACCTGTGGGGAGTGGTCCTGCCTCTGGCGTGGAGCCCACAA
GTTTCCCAGTCGAGTCTGAGGCTCGACTGATGGAGGATGTGCTGAGACCTTTGGAACAGGCATTGGAAGA
CTGCCGTGGCCACACAAGGAAGCAGGTATGTGATGACATCAGCCGACGCCTGGCACTGCTGCAGGAACAG
TGGGCTGGAGGAAAGTTGTCAATACCTGTAAAGAAGAGAATGGCTCTACTGGTGCAAGAGCTTTCAAGCC
ACCGGTGGGACGCAGCAGATGACATCCACCGCTCCCTCATGGTTGACCATGTGACTGAGGTCAGTCAGTG
GATGGTAGGAGTTAAAAGATTAATTGCAGAAAAGAGGAGTCTGTTTTCAGAGGAGGCAGCCAATGAAGAG
AAATCTGCAGCCACAGCTGAGAAGAACCATACCATACCAGGCTTCCAGCAGGCTTCATAATCCTCGGTTC
CCCAGACTCACCGGACACCATCTCCTATGCCTTGGAGACCTTCTGTCACTTGGCTCCCTTCTTACCACCA

CCAAGACTGTCCCACTGGGCCTGACCCACCTATGAGGGAAGAAGTCCCACCTGGGCCAGAGGGAGTTCAT
GTGTTACTCATAACATGCATTTCAATAAAAACATCTCTGCGGTGGAAAAAAAAAAAAAAAAAAA

Human SRA1 mRNA sequence - var7 (public gi: 16549596) (SEQ ID NO: 181)
TTATAGCAAAATCAGTGCAAATAAAAATCCCTCAGTGACCTCACTGGATGTGAGTATATTGGGCCTGGGA
CAGGGCTGGGGGCTAACACCCTGTGTGAGATGAGTGTCTTTGTGTCTGTGCTTGATGTTGGTGGCTCTCT
GTAGTCACATGACAGCATGGGTGTGATGGAGATCTGACTTCATTCAACAAACATATTTTCTAAGGAGTTC
CCTGTGCCAGGCACTAAGCTGGGCACTGGGAATGTAATAAAATAGTCAAGGTCCCACCTTCTAAGACTGT
CCGACAGGGAAACGAACAAGAGTCAAATAAGGCAGAAGATGTGATGTAATACACCTACGAAATCTCAGAG
GGTTGTAGGGTCGTGGGAGCTCAAGTGAGACACTTAACCTGGCCTGAGACATTCCAGAAGGCCTCCTGAA
GAACTGACATCTGAACTGAGAACTGAAGGAAGATGAGTACTAGTGAGGCTACCGGACGTGAATGTGGAGA
TTGTGCAGGGCAATGCAAGAGGAGGCTGTAGAAGTCAACCTGGCTAGATCACAGCGGGGTGTATGTGGGG
CAGGAGCTTCTTTGTTTGAATTTGCTCCTGAGAGGATGAGGCCTCCTAGAGCACTGGCTCCTGGACAGCA
ACCTCCTTTGGTGCCTTGTGACCAGGGCCCTGATGGTTCATTAGATGGAGCCTTCGAGTCTTAGGGAGTT
GCCGCAGGGTCCCCACAGCGGCTCCCGACGGTTGTGAACCAGCATCCATTCTCCACGGATTCCGGCAACC
CGCCTGGCCCTGGACGTGTCTCAACTGGCCCGCGTGAGGGGCCGCCCCGGAAATGACGCGCTGCCCCGCT
GGCCAAGCGGAAGTGGAGATGGCGGAGCTGTACGTGAAGCCGGGCAACAAGGAACGCGGCTGGAACGACC
CGCCGCAGTTCTCATACGGGCTGCAGACCCAGGCCGGCGGACCCAGGCGCTCGCTGCTTACCAAGAGGGT
AGCCGCACCCCAGGATGGATCCCCCAGAGTCCCCGCATCAGAGACTTCTCCTGGGCCTCCCCCAATGGGG
CCTCCACCTCCTTCAAGTAAGGCTCCCAGGTCCCCACCTGTGGGGAGTGGTCCTGCCTCTGGCGTGGAGC
CCACAAGTTTCCCAGTCGAGTCTGAGGCTCGACTGATGGAGGATGTGCTGAGACCTTTGGAACAGGCATT
GGAAGACTGCCGTGGCCACACAAGGAAGCAGGTATGTGATGACATCAGCCGACGCCTGGCACTGCTGCAG
GAACAGTGGGCTGGAGGAAAGTTGTCAATACCTGTAAAGAAGAGAATGGCTCTACTGGTGCAAGAGCTTT
CAAGCCACCGGTGGGACGCAGCAGATGACATCCACCGCTCCCTCATGGTTGACCATGTGACTGAGGTCAG
TCAGTGGATGGTAGGAGTTAAAAGATTAATTGCAGAAAAGAGGAGTCTGTTTTCAGAGGAGGCAGCCAAT
GAAGAGAAATCTGCAGCCACAGCTGAGAAGAACCATACCATACCAGGCTTCCAGCAGGCTTCATAATCCT
CGGTTCCCCAGACTCACCGGACACCATCCCCTATGCCTTGGAGACCTTCTGTCACTTGGCTCCCTTCTTA
CCACCACCAAGACTGTCCCACTGGGCCTGACCCACCTATGAGGGAAGAAGTCCCACCTGGGCCAGAGGGA
GTTCATGTGTTACTCATAACATGCATTTCAATAAAAACATCTCTGCGGTGGGCCTTGGGTAGGAGAGATG
AACCCTTCCGGTGCCAAGCTAGTCCCCTCTGGTGTCCTCGACTGCCCTGCTCCCTGTGTATCTGCAAACC
TCTGTTCTCCCTTCTCCATTCATCAGGAAGGGATCTGCTGGGTAAAGTCAGACTACTGCCTACCACTTTT
TCCCAAAGTAGACTGAAAGCACATCCTGTGCTGGGCGGAGCAGCTGTGTTTGGATGGTTTCATTTCAGCA
TGAGAACAGACTCAAATAGAACGGGGAGACTTTTCCCTCAACAAAAGGAAAGACAGTCCTATTTGCACTG
TATCACCCTTGAGATACTACTGTTACAGAGATTAGAACC

Human SRA1 mRNA sequence - var8 (public gi: 9930609) (SEQ ID NO: 182)
TCCTTTGGTGCCTTGTGACCAGGGCCCTGATGGTTCATTAGATGGAGCCTTCGAGTCTTAGGGAGTTGCC
GCAGGGTCCCCACAGCGGCTCCCGACGGTTGTGAACCAGCATCCATCCTCCACGGATTCCGGCAACCCGC
CTGGCCCTGGACGTGTCTCAACTGGCCCGCGTGAGGGGCCGCCCCGGAAATGACGCGCTGCCCCGCTGGC
CAAGCGGAAGTGGAGATGGCGGAGCTGTACGTGAAGCCGGGCAACAAGGAACGCGGCTGGAACGACCCGC
CGCAGTTCTCATACGGGCTGCAGACCCAGGCCGGCGGACCCAGGCGCTCGCTGCTTACCAAGAGGGTAGC
CGCACCCCAGGATGGATCCCCCAGAGTCCCCGCATCAGAGACTTCTCCTGGGCCTCCCCCAATGGGGCCT
CCACCTCCTTCAAGTAAGGCTCCCAGGTCCCCACCTGTGGGGAGTGGTCCTGCCTCTGGCGTGGAGCCCA
CAAGTTTCCCAGTCGAGTCTGAGGCTGTGATGGAGGATGTGCTGAGACCTTTGGAACAGGCATTGGAAGA
CTGCCGTGGCCACACAAGGAAGCAGGTATGTGATGACATCAGCCGACGCCTGGCACTGCTGCAGGAACAG
TGGGCTGGAGGAAAGTTGTCAATACCTGTAAAGAAGAGAATGGCTCTACTGGTGCAAGAGCTTTCAAGCC
ACCGGTGGGACGCAGCAGATGACATCCACCGCTCCCTCATGGTTGACCATGTGACTGAGGTCAGTCAGTG
GATGGTAGGAGTTAAAAGATTAATTGCAGAAAAGAGGAGTCTGTTTTCAGAGGAGGCAGCCAATGAAGAG
AAATCTGCAGCCACAGCTGAGAAGAACCATACCATACCAGGCTTCCAGCAGGCTTCATAATCCTCGGTTC
CCCAGACT

Human SRA1 protein sequence - var1 (public gi: 9930610) (SEQ ID NO: 291)
MTRCPAGQAEVEMAELYVKPGNKERGWNDPPQFSYGLQTQAGGPRRSLLTKRVAAPQDGSPRVPASETSP
GPPPMGPPPPSSKAPRSPPVGSGPASGVEPTSFPVESEAVMEDVLRPLEQALEDCRGHTRKQVCDDISRR
LALLQEQWAGGKLSIPVKKRMALLVQELSSHRWDAADDIHRSLMVDHVTEVSQWMVGVKRLIAEKRSLFS
EEAANEEKSAATAEKNHTIPGFQQAS

Human SRA1 protein sequence - var2 (public gi: 25123255) (SEQ ID NO: 292)
MGPPPPSSKAPRSPPVGSGPASGVEPTSFPVESEAVMEDVLRPLEQALEDCRGHTRKQVCDDISRRLALL
QEQWAGGKLSIPVKKRMALLVQELSSHRWDAADDIHRSLMVDHVTEVSQWMVGVKRLIAEKRSLFSEEAA
NEEKSAATAEKNHTIPGFQQAS

Human SRA1 protein sequence - var3 (public gi: 9930614) (SEQ ID NO: 293)
MTRCPAGQAEVEMAELYVKPGNKERGWNDPPQFSYGLQTQAGGPRRSLLTKRVAAPQDGSPRVPASETSP
GPPPMGPPPPSSKAPRSPPVGSGPASGVEPTSFPVESEARLMEDVLRPLEQALEDCRGHTRKQVCDDISR
RLALLQEQWAGGKLSIPVKKRMALLVQELSSHRWDAADDIHRSLMVDHVTEVSQWMVGVKRLIAEKRSLF
SEEAANEEKSAATAEKNHTIPGFQQAS

Human SRA1 protein sequence - var4 (public gi: 9930612) (SEQ ID NO: 294)
MTRCPAGQAEVEMAELYVKPGNKERGWNDPPQFSYGLQTQAGGPRRSLLIKRVAAPQDGSPRVPASETSP
GPPPMGPPPPSSKAPRSPPVGSGPASGVEPTSFPVESEAVMEDVLRPLEQALEDCRGHTRKQVCDDISRR
LALLQEQWAGGKLSIPVKKRMALLVQELSSHRWDAADDIHRSLMVDHVTEVSQWMVGVKRLIAEKRSLFS
EEAANEEKSAATAEKNHTIPGFQQAS

Unigene Name: SYNE1 Unigene ID: Hs.416719 Clone ID: 3GD_138aa2938

Human SYNE1 mRNA sequence - var1 (public gi: 21753084) (SEQ ID NO: 183)
GTACAAAAACGAACTTTCACAAAATGGATCAACTCTCATCTGGCCAAGCGGAAACCTCCAATGGTGGTGG
ACGATCTTTTTGAAGACATGAAAGATGGTGTTAAACTGCTTGCCCTTCTGGAGGTCCTGTCTGGCAGAA
ACTGCCTTGTGAACAAGGACGCCGGATGAAGCGAATCCATGCTGTGGCTAACATTGGCACGGCACTCAAG
TTCCTCGAAGGAAGAAAGATTAAATTAGTCAACATTAACTCCACCGATATAGCTGATGGCCGACCCTCAA
TAGTTCTTGGATTGATGTGGACCATTATTCTATATTTCCAGATTGAAGAGTTGACCAGCAACCTGCCCCA
GCTCCAGTCTTTGTCCAGCAGCGCATCCTCCGTGGACAGCATAGTTAGCTCTGAGACTCCCAGCCCACCA
AGTAAACGGAAGGTGACCACCAAGATCCAAGGAAATGCTAAGAAGGCTTTATTAAAGTGGGTTCAGTACA
CAGCTGGCAAGCAGACTGGAATAGAAGTAAAAGATTTTGGGAAGAGTTGGAGAAGCGGGGTTGCCTTTCA
TTCAGTTATTCATGCCATTCGACCGGAATTGGTGGACTTGGAGACAGTGAAAGGCAGATCCAACCGAGAA
AATTTGGAGGATGCTTTCACTATCGCTGAAACAGAACTGGGGATCCCAAGACTGCTAGATCCTGAAGACG
TTGATGTGGATAAACCAGATGAGAAATCTATTATGACCTATGTAGCCCAGTTTCTGAAACATTATCCTGA
CATCCACAATGCAAGCACTGATGGGCAAGAGGATGATGAAATACTTCCAGGTTTCCCATCTTTTGCAAAT
TCTGTACAAAATTTTAAGAGAGAAGACAGAGTAATTTTTAAGGAAATGAAAGTTTGGATAGAACAATTTG
AGAGAGATTTGACAAGAGCACAGATGGTGGAATCAAATTTACAGGATAAATATCAGTCATTTAAGCACTT
CAGAGTTCAATATGAAATGAAGAGGAAACAGATTGAACATTTAATACAACCATTACACAGAGACGGTAAA
TTGTCACTTGACCAAGCATTGGTAAAACAATCTTGGGATAGAGTGACCTCCAGGCTCTTTGACTGGCATA
TACAGCTTGATAAATCTCTTCCTGCACCTCTGGGCACCATAGGTGCCTGGCTGTACAGAGCGGAGGTGGC
CCTGAGAGAGGAAATAACCGTTCAACAGGTCCACGAGGAAACAGCAAACACGATACAACGGAAACTTGAG
CAACATAAGGATCTGCTTCAAAACACGGATGCCCACAAAAGAGCATTCCATGAAATCTACCGGACCAGGT
CTGTTAACGGGATTCCAGTGCCACCTGATCAATTAGAGGACATGGCCGAGAGGTTTCATTTTGTTTCCTC
CACATCAGAGCTACACCTAATGAAAATGGAATTTTTAGAATTAAAGTACCGTCTGCTCTCACTGCTGGTT
CTTGCAGAGTCAAAGCTGAAGTCTTGGATCATTAAGTACGGGAGGAGAGAGTCAGTGGAGCAGCTTCTAC
AAAACTACGTGTCTTTTATAGAAAATAGCAAGTTCTTTGAACAATATGAGGTGACATACCAGATCTTGAA
ACAGACAGCTGAGATGTATGTCAAAGCAGATGGTTCAGTGGAAGAAGCTGAGAATGTGATGAAATTCATG
AATGAAACCACCGCTCAGTGGAGGAATCTCTCAGTAGAAGTGAGGAGTGTGAGGAGCATGCTGGAAGAAG
TGATCTCTAACTGGGATCGCTATGGCAATACAGTGGCTAGTCTGCAAGCCTGGCTAGAGGATGCTGAAAA
AATGCTCAATCAATCAGAAAATGCCAAAAAGGATTTTTTTCGAAATTTACCTCATTGGATTCAGCAGCAT
ACTGCCATGAACGATGCTGGCAATTTTCTAATTGAAACCTGTGATGAGATGGTTTCCCGTGACCTGAAGC
AGCAATTACTGTTGCTAAATGGGCGGTGGAGGGAGTTGTTTATGGAAGTCAAGCAATATGCTCAAGCTGA
TGAGATGGACAGAATGAAGAAGGAATACACAGACTGTGTTGTTACCCTGTCTGCTTTTGCAACGGAAGCC
CATAAGAAACTTTCTGAACCCTTAGAAGTCTCTTTTATGAATGTCAAGCTATTAATTCAAGACTTGGAGG
ATATTGAGCAGAGGGTGCCTGTGATGGATGCCCAATACAAGATAATTACAAAGACAGCACACCTCATTAC
CAAAGAAAGCCCCC

Human SYNE1 mRNA sequence - var2 (public gi: 22382201) (SEQ ID NO: 184)
AGCGGCTGCCTCCTTGTTGAGTGCTGCAAAGGCCTGGAATTCATTTATGACAGAATAGATCTAGAAAAGT
CCAAGCATGTTTTCTAGAGTGGTGTAGCCCTGTGCTGCCTCCAGTGAAGAGTCTCTTGGTGTTGGCTTCG
TGCTTCCGGAGGGACCATGGCAACCTCCAGAGGGGCCTCCCGGTGTCCTCGGGATATCGCCAATGTGATG
CAGAGGCTGCAAGATGAGCAAGAGATAGTACAAAAACGAACTTTCACAAAATGGATCAACTCTCATCTGG
CCAAGCGGAAACCTCCAATGGTGGTGGACGATCTTTTTGAAGACATGAAAGATGGTGTTAAACTGCTTGC
CCTTCTGGAGGTCCTGTCTGGCAGAAACTGCCTTGTGAACAAGGACGCCGGATGAAGCGAATCCATGCT
GTGGCTAACATTGGCACGGCACTCAAGTTCCTCGAAGGAAGAAAGATTAAATTAGTCAACATTAACTCCA
CCGATATAGCTGATGGCCGACCCTCAATAGTTCTTGGATTGATGTGGACCATTATTCTATATTTCCAGAT

TGAAGAGTTGACCAGCAACCTGCCCCAGCTCCAGTCTTTGTCCAGCAGCGCATCCTCCGTGGACAGCATA
GTTAGCTCTGAGACTCCCAGCCAACCAAGTAAACGGAAGGTGACCACCAAGATCCAAGGAAATGCTAAGA
AGGCTTTATTAAAGTGGGTTCAGTACACAGCTGGCAAGCAGACTGGAATAGAAGTAAAAGATTTTGGGAA
GAGTTGGAGAAGCGGGGTTGCCTTTCATTCAGTTATTCATGCCATTCGACCGGAATTGGTGGACTTGGAG
ACAGTGAAAGGCAGATCCAACCGAGAAAATTTGGAGGATGCTTTCACTATCGCCGAAACAGAACTGGGGA
TCCCAAGACTGCTAGATCCTGAAGACGTTGATGTGGATAAACCAGATGAGAAATCTATTATGACATATGT
AGCCCAGTTTCTGAAACATTATCCTGACATCCACAATGCAAGCACTGATGGCAAGAGGATGATGAAATA
CTTCCAGGTTTCCCATCTTTTGCAAATTCTGTACAAAATTTTAAGAGAGAAGACAGAGTAATTTTTAAGG
AAATGAAAGTTTGGATAGAACAATTTGAGAGAGATTTGACAAGAGCACAGATGGTGGAATCAAATTTACA
GGATAAATATCAGTCATTTAAGCACTTCAGAGTTCAATATGAAATGAAGAGGAAACAGATTGAACATTTA
ATACAACCATTACACAGAGACGGTAAATTGTCACTTGACCAAGCATTGGTAAAACAATCTTGGGATAGAG
TGACCTCCAGGCTCTTTGACTGGCATATACAGCTTGATAAATCTCTTCCTGCACCTCTGGGCACCATAGG
TGCCTGGCTGTACAGAGCGGAGGTGGCCCTGAGAGAGGAAATAACCGTTCAACAGGTCCACGAGGAAACA
GCAAACACGATACAACGGAAACTTGAGCAACATAAGGATCTGCTTCAAAACACGGATGCCCACAAAAGAG
CATTCCATGAAATCTACCGGACCAGGTCTGTTAACGGGATTCCAGTGCCACCTGATCAATTAGAGGACAT
GGCCGAGAGGTTTCATTTTGTTTCCTCCACATCAGAGCTACACCTAATGAAAATGGAATTTTTAGAATTA
AAGTACCGTCTGCTATCACTGCTGTTCTTGCAGAGTCAAAGCTGAAGTCTTGGATCATTAAGTACGGGAG
GAGAGAGTCAGTGGAGCAGCTTCTACAAAACTACGTGTCTTTTATAGAAAATAGCAAGTTCTTTGAACAA
TATGAGGTGACATACCAGATCTTGAAACAGACAGCTGAGATGTATGTCAAAGCAGATGGTTCAGTGGAAG
AAGCTGAGAATGTGATGAAATTCATGAATGAAACCACCGCTCAGTGGAGGAATCTCTCAGTAGAAGTGAG
GAGTGTGAGGAGCATGCTGGAAGAAGTGATCTCTAACTGGGATCGCTATGGCAATACAGTGGCTAGTCTG
CAAGCCTGGCTAGAGGATGCTGAAAAAATGCTCAATCAATCAGAAAATGCCAAAAAGGATTTTTTTCGAA
ATTTACCTCATTGGATTCAGCAGCATACTGCCATGAACGATGCTGGCAATTTTCTAATTGAAACCTGTGA
TGAGATGGTTTCCCGTGACCTGAAGCAGCAATTACTGTTGCTAAATGGGCGGTGGAGGGAGTTGTTTATG
GAAGTCAAGCAATATGCTCAAGCTGATGAGATGGACAGAATGAAGAAGGAATACACAGACTGTGTTGTTA
CCCTGTCTGCTTTTGCAACGGAAGCCCATAAGAAACTTTCTGAACCCTTAGAAGTCTCTTTTATGAATGT
CAAGCTATTAATTCAAGACTTGGAGGTGAGGGGTTTCTGAATCAAAATGAAAAGCCTACTCTGTTGTGAG
AGGAAAGATCAGCAAGTTTATTCAGATCATTGCAAAAGCTGTTTCTGTGTCTCCTGGGCATCATTTTGAC
ATGTCTGTATGTCCCAATTTGCACCTGTCAGAAAAAAATGTATTGAACATAAAAAAGACATGACTTGATC
ATATAAAGTAACTTCAAATTGTTAAAAAAAAAAAAAAAGAAAAAAAAAAAAA

Human SYNE1 mRNA sequence - var3 (public gi: 28192627) (SEQ ID NO: 185)
AGTACGCGGGAGTCTTAAAACGGAAGAAGAAAAAGAAGCAGTTCAGTCTTTGGGAGAGCTGCCTCCTTGT
TGAGTGCTGCAAAGGCCTGGAATTCATTTATGACAGAATAGATCTAGAAAAGTCCAAGCATGTTTTCTAG
AGTGGTGTAGCCCTGTGCTGCCTCCAGTGAAGAGTCTCTTGGTGTTGGCTTCGTGCTTCCGGAGGGACCA
TGGCAACCTCCAGAGGGGCCTCCCGGTGTCCTCGGGATATCGCCAATGTGATGCAGAGGCTGCAAGATGA
GCAAGAGATAGTACAAAAACGAACTTTCACAAAATGGATCAACTCTCATCTGGCCAAGCGGAAACCTCCA
ATGGTGGTGGACGATCTTTTTGAAGACATGAAAGATGGTGTTAAACTGCTTGCCCTTCTGGAGGTCCTGT
CTGGGCAGAAACTGCCTTGTGAACAAGGACGCCGGATGAAGCGAATCCATGCTGTGGCTAACATTGGCAC
GGCACTCAAGTTCCTCGAAGGAAGAAAGATTAAATTAGTCAACATTAACTCCACCGATATAGCTGACGGC
CGACCCTCAATAGTTCTTGGATTGATGTGGACCATTATTCTATATTTCCAGATTGAAGAGTTGACCAGCA
ACCTGCCCCAGCTCCAGTCTTTGTCCAGCAGCGCATCCTCCGTGGACAGCATAGTTAGCTCTGAGACTCC
CAGCCCACCAAGTAAACGGAAGGTGACCACCAAGATCCAAGGAAATGCTAAGAAGGCTTTATTAAAGTGG
GTTCAGTACACAGCTGGCAAGCAGACTGGAATAGAAGTAAAAGATTTTGGGAAGAGTTGGAGAAGCGGGG
TTGCCTTTCATTCAGTTATTCATGCCATTCGACCGGAATTGGTGGACTTGGAGACAGTGAAAGGCAGATC
CAACCGAGAAAATTTGGAGGATGCTTTCACTATCGCCGAAACAGAACTGGGGATCCCAAGACTGCTAGAT
CCTGAAGACGTTGATGTGGATAAACCAGATGAGAAATCTATTATGACCTATGTAGCCCAGTTTCTGAAAC
ATTATCCTGACATCCACAATGCAAGCACTGATGGCAAGAGGATGATGAAATACTTCCAGGTTTCCCATC
TTTTGCAAATTCTGTACAAAATTTTAAGAGAGAAGACAGAGTAATTTTTAAGGAAATGAAAGTTTGGATA
GAACAATTTGAGAGAGATTTGACAAGAGCACAGATGGTGGAATCAAATTTACAGGATAAATATCAGTCAT
TTAAGCACTTCAGAGTTCAATATGAAATGAAGAGGAAACAGATTGAACATTTAATACAACCATTACACAG
AGACGGTAAATTGTCACTTGACCAAGCATTGGTAAAACAATCTTGGGATAGAGTGACCTCCAGGCTCTTT
GACTGGCATATACAGCTTGATAAATCTCTTCCTGCACCTCTGGGCACCATAGGTGCCTGGCTGTACAGAG
CGGAGGTGGCCCTGAGAGAGGAAATAACCGTTCAACAGGTCCACGAGGAAACAGCAAACACGATACAACG
GAAACTTGAGCAACATAAG

Human SYNE1 mRNA sequence - var4 (public gi: 21734187) (SEQ ID NO: 186)
GGGACACAGTGAGAAACCACTGTATGAAGTGTTTGGGGCTCAGATGGCAGTGTGGTGGCAGGACACACAA
GCAGGGGCAACAAAGCCGGAGTCCTGGGAGAGACTTCGGGAAAGAAACCATGGAGCAGAGTCCAGAGGGT
GAATCGCAGTTGGTCAGGTGGGCTCGGAAGTGCTTAGAGGAAGAAAGGAGGGGAAGCCTCCCTGCTTTGT
GATTGGCCTGGTGCTTTTGATCATTGGCAGTTTGTCTGGAATGACAGGCATGGAATGGACAAAGTGGAGA
AGAGCCTGGCTGTGAAGCAGCCTGTTCACCATGGAAAAGAGCTGTGGCTCCTATCCTGAAGGTCGTGGAG

CCACAGCAGGATCTGCGGAGGGAGGTGCTGGGATCCTCCCTCCTCAGGGATGTGCAGATTTTCATATTGT
ATCTTTCTGGATACCACAGGGAGAAGGGCATATTCGGCGGAGAGAGACCAAATGAAACCTTTTACAACCT
CAGACAGAAGTAGGGTGGTGGCCTAAACTAGGGGAAGCAGAATTGGGAATGGGGAGAATGGGAATGATGT
GAGAAATCACATAGAGAAGACTCCTCCAGAACTCTCAGTCCATTGAACTGGGATGGAGGCGATTTTCTGG
GCTGGGCATCTTGGTGAAAGATGCAGGTGGTCCTAGGCCCTGAGGACCACAAGAGGGAAGGAGCACTGTG
GGTGCAGGTGGGCAAGGGAGGTGGGGCTGTGAGAGCAGGGAGGGGATGAGTTTGCTTGTGTGCATCCTGA
TCTTGAGATACCTGCAGAATATCCAAATGCAAAAGTCCAGTCGTAGATGCACGGTGTGAAGTGCAGAAGC
CAGAAATGCAGATTGGGTAGGTATTCACATGTAAATGGCAATGGTCCTGGAGTGAACGGAGGAGCTCCCA
CAGGAAGAGTGTGTGAAGGAAAACAAGAAGGACCACCACCCAAGCCACACATGCAGTGAAGGGATGGACA
GAGAAACAGAAACTCTGTAAGGAAGGTGAATAAAAATAGAATAAAGAGTTGGAGGCTGATTTGTGGCACT
TGGAAATGTATCTCATACATTCTGTCAAAGGACATCTGGGGAATTTCTGTTTGGTTCTGGTGGTTCACAT
CAGATTCCCAAGGGATGACACTGTTCTAAAAAGAAAATGATTTCTCTCATTTCTATTTTGTCTTTACAGT
AAGGCCTATTAGTCAGGCATATGGCATCTGAAGCAGAGCTGTCCAAAACCAGCCACTGGCCAGTTGGGAC
TGTTGAGCTCTGAGATGGGACTGTGCAAATTGAGATGGGTTGTGCGTGGAAAACATGCTTACATGAATTT
CAAAGACTTAGTACAAGAAAGAAAATAAAATATTAATAATTATATTGATTACATGTTATAATCCCTGTCT
AATGTAGTGTTAAAATTAATTTTATAAGTTTCTTTTTACATTTCTAATGTGGCTACGAAACCTTTAAGAT
TACATATATAGTTCACATAGAAATATATGGGACAGCGCTGCTCTGGAGTCTGGGCTGAAATCTCAGTTCT
GCCATGTACTTTCTGTTTAAACTTAGATAAGGAACCTAATTCCTCTGTGCCTCAGTTTTCTCATCTATAA
AATGGGAATAACATTCCCAGGTACCCTATAGGGTTTCTATGTGATAAATTTGTGCTCAGACCAGAGCCTG
GCTCATAAAAACACTCTCAGTCACTGTGAGTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGA
AGTCTAGTTCTGTTGCCCAGGCTGGAGTGCAGTGGCACAATCTCGGCTCACTGCAACCTCCGCCTCCCCG
GTTCAAGCGATCCTCCTGCCTCAGCCTCCCATGTAGCTGGGATTACAGGTACCTGTCACCACACCTGGCT
AATTTTTGTATTTTTAGTAGAGATGGGGTTCCACCATGTCGGCCAGGCTGGTCTCGAATTCCTGACCTCA
GGTGATCCACCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCCACCCACTG
TGAATTCTTATGATTCATTTCAGGAAAGCTTTGGTGAGCCTGCACGCTCCTCTGTGCGCTCAGGAGCTAT
GTGTCTAGAATAACTGACTTCCTTTTTTTTCCCCTAGGAAAGTTATTTTCCTGCACAAGGGATTCAGGGT
TTCCAGAACTTAGCTGTCAACTTAGACTGTGCTTTTTTTGCAGATGTAATGATCCCCGAGAGCCCTGAGG
CCTATGTAAAACTCACAGAAAATGCAATCAAAAATACCTCCGGTAGGCACAACAAGCCGGAGCCAGCCTC
CCTCCCTCCCCCGTAGAGCAATCCTCCTCTCTTAGCACATCTGCTGTCTCTCTCCCAGCTTTGTGGCTCT
AGCATGTTAAGGCACAGCCTTCCTCTCTTACTGCTGTACTAGAAAAAACAGCTGGTTAAATCCACACCGA
GAATAAGATTTCACTAATCGAGCGAAATAAAATAACTTCTCAACTTGTTAAATGGTGATTGGTCTCATTA
GGTATAGACCTCTCATGTCCATTAACTGCAGAAAATATGAGAAGGAAAACCCAGTCATCAGCCTCTGCGC
CCTAGTGTTCTACGTGGTGTGGTAATTTCAGCTTCACTGCATGCAGACCTACCTGTGGCTGGAGACTCAG
GGTGCAGGCTCTGGTCCCAAGTCCGCCAGCCTGCATGAGTGACCTTTGGCCCATCCACCACTTTATCCTC
CTCATCTCAAGAATCCCGTATGAGACAAGGGGTGAGATCAGATTTCAGCTCTAAAAAATATATGTAATTT
TAATTTAAGAGGTCGTAAAAGATAATTTGAAATGAAAAATGTATTTCACGGTATGCTGAGCCATAGATAA
GAACAGAACTATTCCTGAAACAAGAAGAATTAAAGAAAGAAAAATGGAACTAGTTTTGCTTAGTGTTGTG
TTAGACAAACTGTAGTGCAGGAGTTCAGGATCAGTGATGTGGGATGTCGGCGGGGAATAGAGTTTGAACG
CAGTGATATGATATTGAATCACGGAGTTACTAGTTTACGCTTCAGTTTTTGAAGAAAATCAAAAGGACAG
AAAGCAAAGTAACATTACTGAGAGGGTGATTCCCAGGGAGGGACCTCTCCTAGGTGTATCTAGAAGGCCT
TTTTTTAGAAACAAATAAAAACATTTAATAAAGCTTACTAATATTTGTTCTGCTTTCACCCCATGCTAGC
TTCACTGATGATCAAATGTTCCTGTGTAGTTCGAAGACTTTGACACACACACACACACACACACACACAC
TCAGTAATTTTACAAAGAAATGTTACAACTTTGAGAGGAGAATGAGCCAGAATTCTAGGCTATGAGTAAG
AACCTGCCTAGATGGAAATGTTAAAATCTTAGCTTTCTCCTGGTTTTGTTTCAGATCTTAGATAAAAAGC
AAGTCGTTGCTAGTTTGATATCTCTGTATATATCTATTCTGAGGCACTCTTTTCTTGATTAATGAATTTA
TGCCTTCAGTAAATGATGCAGCAACCTGAGCCTTCCGTGACACTATCTTCCCCTGAGGTGCATGAAGAAA
AATCAGAGGGAGGATCTTCCCCTGCTCACTAAGCGATAGCAGAAAGAACATGAGAAAAAGAACAGCTTTC
TCCTTACTGAGATGCAGTAGACACCATTGCAGATTTTTAGAAAGGCCTTCCACACTGACAACATTCAGAT
GATCAGGGGTTCAGCTGGAAGAGGGTCAGGTGCACGAAGCTGTTCAAAACGGACTGGAGAGCCGTTTTGC
GACGTCCGATCTGCTAGGGCTCTCAGTCAAGCACCTTCATTGGTTTGGCCATTTTAATGTCTAACCCACC
GGCACGATGGTACTAAATTTGCTTTGTAATTACCCACACCTGCCATTTCTATGCTGCTGTAACTGAAATC
ATTTCTGAAACTTCTCCTTGAATATCAAGCTTTAAATAAGTCAAATAGTTTGTCCAGTAAAGATTCTTAT
GGTTGCCACCGCAGGGGACCACAGTGCCCTAGAGTCACAGATCCGACAACTGGGCAAAGCCCTGGATGAT
AGCCGCTTTCAGATACAGCAAACCGAAAATATCATTCGCAGCAAAACTCCCACGGGGCCGGAGCTAGACA
CCAGCTACAAAGGCTACATGAAACTGCTGGGCGAATGCAGTAGCAGTATAGACTCCGTGAAGAGACTGGA
GCACAAACTGAAGGAGGAAGAGGAGAGCCTTCCTGGCTTTGTTAACCTGCATAGTACCGAAACCCAAACG
GCTGGTGTGATTGACCGATGGGAGCTTCTCCAGGCCCAGGCATTGAGCAAGGAGTTGAGGATGAAGCAGA
ACCTCCAGAAGTGGCAGCAGTTTAACTCAGACTTGAACAGCATCTGGGCCTGGCTGGGGGACACGGAGGA
GGAGTTGGAACAGCTCCAGCGTCTGGAACTCAGCACTGACATCCAGACCATCGAGCTCCAGATCAAAAAG
CTCAAGGAGCTCCAGAAAGCTGTGGACCACCGCAAAGCCATCATCCTCTCCATCAATCTCTGCAGCCCTG
AGTTCACCCGGGCTGACAGCAAGGAGAGCCGGGACCTGCAGGATCGCTTGTCGCAGATGAATGGGCGCTG
GGACCGAGTGTGCTCTCTGCTGGAGGAGTGGCGGGGCCTGCTGCAGGATGCCCTGATGCAGTGCCAGGGT
TTCCATGAAATGAGCCATGGTTTGCTTCTTATGCTGGAGAACATTGACAGAAGGAAAAATGAAATTGTCC

CTATTGATTCTAACCTTGATGCAGAGATACTTCAGGACCATCACAAACAGCTTATGGTAAGATGTGTGAA
CTCTGGCAGCCTCCAGTTATTTTTAGCAGGGTTGCATTTCATTTACAGAAAATGAATATAAGTGGTAAGT
GTTGTTTCTTTTTTTTTAACTTTTTGCATTATAGTCTCTACTTTACACTTTTTTAACTCCCTGTGGTTTC
CAATCTTTGTAAAGCAAACATGTGCATAGAAGATGATATCTGCTAGCTTTAGAATCTGATTCTAAAGTTG
TTGCTCAGTTGTAAAAATCTTAGTGTTCTCAAGCAATCTTAATTAGCTTGTGTTGTTTATTAAGGCAGCT
TAATTTAAACTTCATGTTACATCTATGGCCCAAAAGTATATTTGGTGGCTTGTAGTAAAAGGTCATTAAA
ATATTAGAATAGAATGAGACAATTAAGTCTTTTGTTTGTTTTTGTTGTCGTTGTTTTTGAGACAGAGTC
TCACTCTGTTGCCCAGGGTGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAATCTCCGCCTCCGGGATTC
GAGCAATTCTCCTGCCAATTCTCCTGCCTCACCCTCCCGAGTAGGTGGGACTACAGGTGTAAGCCACCAC
GCCTCGCTAATGTTTGTATTTTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCC
TGGCCGCAGGTGATCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCACACCCA
GCCGAGTCTTTCAAAGAGGAATTAAATACATCAGATTAAACATGAACCTGAGCATCAAGTTTTCTGAAAG
CCAAGACAAAATGGGAAACAAGGAGTAAACTTACTTTCATTATCTGGCAAAAACAAAACATACCACTTCT
CAAGGAAGGAGAAACTTTTTCTAGCACTAAATTCAAGAGGAAATTAACTGGTAGACTCTTATACAAGGAT
CTTTGGACAATATAATGTACAGTATATTTAAGTGACTTTATAGAAGATAAGGAAGCATATTTGAGTTCCA
TTAGAAGAAAATATTATGCACTTTGTAGCTCTCTGTATTTTTAAAATGTTATGTCTTAACATTTAACACT
CACCTAAACTACAGAATTGGTACCTTTTAATTCAGTACCATAATAGTCTTAGAAACCTAGAGGAAATAGC
TGTGGAACTGCATTTTTACTTCACTTTGACCTCTGGCATCAAGCTGTGAATGACGAATCACCCCTTTTTT
TTTTCAAATCTTGACTAGATATCAGAGGATACCTAGACATACTTCTGCTTCGCTATATTTAATGTTGTGC
TTTTCTGTTTAAAAGATTATCTTACATCTCACTTGCATACTAATCTATATTTTAATTACTGTCATATATA
CATTAACTAATTTGAACCTTCCAATAATACTGTGGACCAGGCATCAAATCAAACTGAGATCAGAGACGGT
CAGGGGTCTTATAGAATATTTTTGGCAGAGGCAGGATTAGAACTCAGGCCGTGAGCTGCTGCATCCTTTA
GTGTGTGAGCTCCACGTTTGATGCTCAGGTATAATTTCCCACCAAGTTAAGTTGATTGCATTCTGCACTT
TTGGAGCTTTTGCCAATTATCAAAAATGCTTAGAAAAATTAATTTGTTTTTGTATGCATAGCAAATAAAG
CATGAGCTGTTGGAATCCCAACTCAGAGTAGCCTCTTTGCAAGACATGTCTTGCCAACTACTGGTGAATG
CTGAAGGAACAGACTGTTTAGAAGCCAAAGAAAAAGTCCATGTTATTGGAAATCGGCTCAAACTTCTCTT
GAAGGAGGTCAGTCGTCATATCAAGGAACTGGAGAAGTTATTAGACGTGTCAAGTAGTCAGCAGGATTTG
TCTTCCTGGTCTTCTGCTGATGAACTGGACACCTCAGGGTCTGTGAGTCCCACATCAGGAAGGAGCACCC
CAAACAGACAGAAAACGCCACGAGGCAAGTGTAGTCTCTCACAGCCCGGACCCTCTGTCAGCAGTCCACA
TAGCAGGTCCACAAAAGGTGGCTCCGATTCCTCCCTTTCTGAGCCAGGGCCAGGTCGGTCCGGCCGCGGC
TTCCTGTTCAGAGTCCTCCGAGCAGCTCTTCCCCTTCAGCTTCTCCTGCTCCTCCTCATCGGGCTTGCCT
GCCTTGTACCGATGTCAGAGGAAGACTACAGCTGTGCCCTCTCCAACAACTTTGCCCGGTCATTCCACCC
CATGCTCAGATACACGAATGGCCCTCCTCCACTCTGAACTAAGCAGATGCCATCTGCAGAAGTGCTGGTA
GCATAAGGAGGATCGGGTCATAAGCAATCCCAAACTACCAACAAGAGGACCTTGATCTTGGCGAAAGCCC
TCGGTGTGGCAGCTTTAGCCCTCCTCCAGATCACATGTGTGCAAATTATGGCTTCAGAGGTGGAAGATAA
ACAGTGACGGGGGAACAAACAGACAACAAGAAGGTTTGGAAGAAATCTGGTTTGAGACTCTGAACCTTAG
CACTAAGGAGATTGAGTAAGGACCTCCAAAGTTCCCCGGACTCATGAATTCTGGGCCCTTGGCCCATTCT
GTGCACAGCCAAGGACTTCAGTAGACCATCTGGGCAGCTTTCCCATGGTGCTGCTCCAACCATCAGATAA
ATGACCCTCCCAAGCACCATGTCAGTGTCGTACAATCTACCAACCAACCAGTGCTGAAGAGATTTTAGAA
CCTTGTAACATACAATTTTTAAGAGCTTATATGGCAGCTTCCTTTTTACCTTGTTTTCCTTTGGGGCATG
ATGTTTTAACCTTTGCTTTAGAAGCACAAGCTGTAAATCTAAAAGGCACTTTTTTTTAGAGGTATAAAGA
AAAACTAGACGTAATAAATAAGATCATGGAAGGCTTTATGTGAAAAAAGTTGAATGTTATAGTAAAAAAA
AAAGATATTTATGTATGTACAGTTTGCTAAAGCCAAGTTTTGTTTGTATTGATTTCTTTGCATTTATTAT
AGATATTATAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human SYNE1 mRNA sequence - var5 (public gi: 21734305) (SEQ ID NO: 187)
CACTGGCACGACCGCTCTGCAGACAGCCTGCTTTCTCCACAGCCTTCCTCCAATCTCTCCCTCTCGCTCG
CTCAGCCCCTCCGGAGCGAGCGGTCAGGACGAGACACCCCAGCTAGTGTGGACTCCATCCCCCTGGAGTG
GGATCACGACTATGACCTCAGTCGGGACCTGGAGTCTGCAATGTCCAGAGCTCTGCCCTCTGAGGATGAA
GAAGGTCAGGATGACAAAGATTTCTACCTCCGGGGAGCTGTTGCCTTATCAGATGTAATGATCCCCGAGA
GCCCTGAGGCCTATGTAAAACTCACAGAAAATGCAATCAAAAATACCTCCGGGGACCACAGTGCCCTAGA
GTCACAGATCCGACAACTGGGCAAAGCCCTGGATGATAGCCGCTTTCAGATACAGCAAACCGAAAATATC
ATTCGCAGCAAAACTCCCACGGGGCCGGAGCTAGACACCAGCTACAAAGGCTACATGAAACTGCTGGGCG
AATGCAGTAGCAGTATAGACTCCGTGAAGAGACTGGAGCACAAACTGAAGGAGGAAGAGGAGAGCCTTCC
TGGCTTTGTTAACCTGCATAGTACCGAAACCCAAACGGCTGGTGTGATTGACCGATGGGAGCTTCTCCAG
GCCCAGGCATTGAGCAAGGAGTTGAGGATGAAGCAGAACCTCCAGAAGTGGCAGCAGTTTAACTCAGACT
TGAACAGCATCTGGGCCTGGCTGGGGGACACGGAGGAGGAGTTGGAACAGCTCCAGCGTCTGGAACTCAG
CACTGACATCCAGACCATCGAGCTCCAGATCAAAAAGCTCAAGGAGCTCCAGAAAGCTGTGGACCACCGC
AAAGCCATCATCCTCTCCATCAATCTCTGCAGCCCTGAGTTCACCCAGGCTGACAGCAAGGAGAGCCGGG
ACCTGCAGGATCGCTTGTCGCAGATGAATGGGCGCTGGGACCGAGTGTGCTCTCTGCTGGAGGAGTGGCG
GGGCCTGCTGCAGGATGCCCTGATGCAGTGCCAGGGTTTCCATGAAATGAGCCATGGTTTGCTTCTTATG
CTGGAGAACATTGACAGAAGGAAAAATGAAATTGTCCCTATTGATTCTAACCTTGATGCAGAGATACTTC
AGGACCATCACAAACAGCTTATGCAAATAAAGCATGAGCTGTTGGAATCCCAACTCAGAGTAGCCTCTTT

GCAAGACATGTCTTGCCAACTACTGGTGAATGCTGAAGGAACAGACTGTTTAGAAGCCAAAGAAAAAGTC
CATGTTATTGGAAATCGGCTCAAACTTCTCTTGAAGGAGGTCAGTCGTCATATCAAGGAACTGGAGAAGT
TATTAGACGTGTCAAGTAGTCAGCAGGATTTGTCTTCCTGGTCTTCTGCTGATGAACTGGACACCTCAGG
GTCTGTGAGTCCCACATCAGGAAGGAGCACCCCAAACAGACAGAAAACGCCACGAGGCAAGTGTAGTCTC
TCACAGCCTGGACCCTCTGTCAGCAGTCCACATAGCAGGTCCACAAAAGGTGGCTCCGATTCCTCCCTTT
CTGAGCCAGGGCCAGGTCGGTCCGGCCGCGGCTTCCTGTTCAGAGTCCTCCGAGCAGCTCTTCCCCTTCA
GCTTCTCCTGCTCCTCCTCATCGGGCTTGCCTGCCTTGTACCAATGTCAGAGGAAGACTACAGCTGTGCC
CTCTCCAACAACTTTGCCCGGTCATTCCACCCCATGCTCAGATACACGAATGGCCCTCCTCCACTCTGAA
CTAAGCAGATGCCATCTGCAGAAGTGCTGGTAGCATAAGGAGGATCGGGTCATAAGCAATCCCAAACTAC
CAACAAGAGGACCTTGATCTTGGCGAAAGCCATCGGTGTGGCAGCTTTAGCCCTCCTCCAGATCACATGT
GTGCAAATTATGGCTTCAGAGGTGGAAGATAAACAGTGACGGGGAACAAACAGACAACAAGAAGGTTTG
GAAGAAATCTGGTTTGAGACTCTGAACCTTAGCACTAAGGAGATTGAGTAAGGACCTCCAAAGTTCCCCG
GACTCATGAATTCTGGGCCCTTGGCCCATTCTGTGCACAGCCAAGGACTTCAGTAGACCATCTGGGCAGC
TTTCCCATGGTGCTGCTCCAACCATCAGATAAATGACCCTCCCAAGCACCATGTCAGTGTCGTACAATCT
ACCAACCAACCAGTGCTGAAGAGATTTTAGAACCTTGTAACATACAATTTTTAAGAGCTTATATGGCAGC
TTCCTTTTTACCTTGTTTTCCTTTGGGGCATGATGTTTTAACCTTTGCTTTAGAAGCACAAGCTGTAAAT
CTAAAAGGCACTTTTTTTTAGAGGTATAAAGAAAAACTAGATGTAATAAATAAGATCATGGAAGGCTTTA
TGTGAAAAAAGTTGAATGTTATAGTAAAAAAAAAAAGATATTTATGTATGTACAGTTTGCTAAAGCCAAG
TTTTGTTTGTATTGATTTCTTTGCATTTATTATAGATATTATAAAATAAAAAAAAAAAAAAAAAAAAA

Human SYNE1 mRNA sequence - var6 (public gi: 21750070) (SEQ ID NO: 188)
TCAGAGGGTGCTCAATGCTTTCCTGAAAGCTTGTGATGAACTCACCGACATCCTTCCAGAGCAGGAGCAG
CAGGGGCTGCAGGAAGCTGTTCGAAAGCTCCACAAACAATGGAAGGATCTTCAAGGAGAAGCCCCCTTATC
ATTTGCTTCATCTGAAGATTGATGTGGAGAAGAATAGGTTCTTAGCCTCTGCAGAAGAATGCAGAACTGA
GCTGGATCGAGAGACCAAGCTGATGCCCCAGGAAGGCAGTGAAAAGATAATTAAAGAGCACAGGGTTTTC
TTCAGTGACAAAGGTCCTCATCATCTCTGTGAGAAAAGGTTACAGCTCATCGAGGAACTCTGTGTGAAAC
TCCCAGTGCGGGACCCAGTAAGGGACACACCTGGAACCTGTCACGTGACTCTCAAAGAGCTCAGAGCTGC
CATTGACAGCACCTACAGGAAGCTCATGGAAGACCCAGACAAGTGGAAGGACTACACTAGCAGATTCTCT
GAGTTCTCATCTTGGATATCTACAAATGAGACACAATTAAAGGGGATCAAGGGTGAGGCCATCGATACTG
CCAACCACGGAGAGGTTAAACGTGCCGTTGAAGAGATCAGAAATGGTGTTACCAAAAGGGGTGAGACCCT
CAGCTGGCTGAAATCCAGGCTGAAAGTTTTGACAGAAGTTTCTTCTGAGAATGAAGCCCAAAAGCAGGGA
GATGAGCTGGCAAAATTATCCAGCTCTTTCAAGGCTCTTGTGACGCTGCTGTCAGAGGTTGAAAAGATGC
TAAGCAATTTTGGGGACTGTGTCCAGTACAAAGAAATAGTCAAAAATTCTCTCGAAGAATTAATTTCTGG
CTCTAAAGAAGTCCAGGAACAAGCTGAGAAGATCTTGGATACTGAAAATCTGTTTGAAGCACAGCAGTTA
CTTCTTCATCACCAGCAAAAGACAAAGCGGATCTCAGCAAAGAAGAGAGATGTGCAGCAGCAGATCGCGC
AGGCGCAGCAGGGAGAAGGGGGGCTGCCTGACCGAGGCCACGAGGAGCTGCGGAAGCTGGAGAGCACACT
GGATGGCCTGGAGCGCAGCCGGGAGAGGCAGGAACGCCGCATCCAGGTCACATTAAGAAAATGGGAGCGA
TTTGAAACAAACAAAGAAACAGTAGTAAGATACCTTTTTCAAACAGGTTCCAGTCATGAACGCTTCTTGA
GTTTTAGCAGTTTGGAAAGTTTATCTTCAGAACTGGAACAAACAAAGGAGTTTTCTAAACGGACAGAAAG
TATTGCAGTCCAGGCTGAGAACCTTGTAAAGGAAGCTTCAGAGATACCGCTTGGGCCCCAAAATAAGCAG
CTGCTTCAACAGCAGGCCAAGTCAATCAAAGAACAAGTCAAAAAATTAGAAGACACGCTTGAAGAAGAGT
ATGTGATTGACAAGTCCTAAACTTTCTTCTCTGAGATAAAGTTTCATACAATCTTTCCTGTACCTTGTAT
TCAAAACACTCTTAAAATCTCAAAGTGTCTGTGTATTTCAGCATGTTTTGAGGAAACAACTCACAGTTCA
AAAGAAAGTATCGCTAATACAGAAACCAATATCTATAACAGAGCCCAAAAAATATAAAGGATGTGGGTTT
TGCATCTTAAACTGATCATGTTCATGAGAAAGCCATATCTATTCTATTCTGTGGCCTTTGTACATTGTAG
AGGGAATCTTGAAAAAGAACTAATATTTAAAATAATTTTTTTACTATATTATTCTGCTGTCACCATTTAG
AGCGAAAAGGAGATATTTTGTTAGTGTAGATTCCAGGCCTAAATACACATCACATAGACCATATATCTCC
AACCTGAAGAAGCTCCTGGAGCTTGTTTACAGTGCCTCGGTATTCAAGTTATCCTGACTAATATGCTCTT
TCCAGAAATTAACTTTAAAATATTTTATTTTTAAACTTTTAATGTTTGTTTATCTG

Human SYNE1 mRNA sequence - var7 (public gi: 28192521) (SEQ ID NO: 189)
CATATACAGCTTGATAAATCTCTTCCTGCACCTCTGGGCACCATAGGTGCCTGGCTGTACAGAGCGGAGG
TGGCCCTGAGAGAGGAAATAACCGTTCAACAGGTCCACGAGGAAACAGCAAACACGATACAACGGAAACT
TGAGCAACATAAGAGAAAATGCCGGACAATGATGGATCTGCTTCAAAACACGGATGCCCCACAAAAGAGCA
TTCCATGAAATCTACCGGACCAGGTCTGTTAACGGGATTCCAGTGCCACCTGATCAATTAGAGGACATGG
CCGAGAGGTTTCATTTTGTTTCCCCCACATCAGAGCTACACCTAATGAAAATGGAATTTTTAGAATTAAA
GTACCGTCTGCTCTCACTGCTGGTTCTTGCAGAGTCAAAGCTGAAGTCTTGGATCATTAAGTACGGGAGG
AGAGAGTCAGTGGAGCAGCTTCTACAAAACTACGTGTCTTTTATAGAAAATAGCAAGTTCTTTGAACAAT
ATGAGGTGACATACCAGATCTTGAAACAGACAGCTGAGATGTATGTCAAAGCAGATGGTTCAGTGGAAGA
AGCTGAGAATGTGATGAAATTCATGAATGAAACCACCGCTCAGTGGAGGAATCTCTCAGTAGAAGTGAGG
AGTGTGAGGAGCATGCTGGAAGAAGTGATCTCTAACTGGGATCGCTATGGCAATACAGTGGCTAGTCTGC
AAGCCTGGCTAGAGGATGCTGAAAAAATGCTCAATCAATCAGAAAATGCCAAAAAGGATTTTTTTCGAAA

TTTACCTCATTGGATTCAGCAGCATACTGCCATGAACGATGCTGGCAATTTTCTAATTGAAACCTGTGAT
GAGATGGTTTCCCGTGACCTGAAGCAGCAATTACTGTTGCTAAATGGGCGGTGGAGGGAGTTGTTTATGG
AAGTCAAGCAATATGCTCAAGCTGATGAGATGGACAGAATGAAGAAGGAATACACAGACTGTGTTGTTAC
CCTGTCTGCTTTTGCAACGGAAGCCCATAAGAAACTTTCTGAACCCTTAGAAGTCTCTTTTATGAATGTC
AAGCTATTAATTCAAGACTTGGAGGATATTGAGCAGAGGGTGCCTGTGATGGATGCCCAATACAAGATAA
TTACAAAGACAGCACACCTCATTACCAAAGAAAGCCCCCAAGAAGAAGGAAAAGAAATGTTTGCGACCAT
GTCAAAGCTCAAAGAGCAGCTAACCAAGGTCAAAGAATGTTACTCCCCACTCCTTTATGAGTCTCAGCAG
CTGTTGATTCCGTTGGAGGAATTAGAAAAGCAGATGACGTCCTTTTATGACTCACTTGGGAAAATCAATG
AAATTATCACAGTTCTTGAGCGTGAGGCACAATCGAGTGCCCTTTTAAACAAAAACATCAGG

Human SYNE1 mRNA sequence - var8 (public gi: 19584384) (SEQ ID NO: 190)
AAGCTATTAATTCAAGACTTGGAGGATATTGAGCAGAGGGTGCCTGTGATGGATGCCCAATACAAGATAA
TTACAAAGACAGCACACCTCATTACCAAAGAAAGCCCCCAAGAAGAAGGAAAAGAAATGTTTGCGACCAT
GTCAAAGCTCAAAGAGCAGCTAACCAAGGTCAAAGAATGTTACTCCCCACTCCTTTATGAGTCTCAGCAG
CTGTTGATTCCGTTGGAGGAATTAGAAAAGCAGATGACGTCCTTTTATGACTCACTTGGGAAAATCAATG
AAATTATCACAGTTCTTGAGCGTGAGGCACAATCGAGTGCCCTTTTAAACAAAAACATCAGGAACTGTT
AGCTTGTCAAGAAAACTGTAAGAAAACCTTGACACTTATTGAGAAAGGCAGTCAAAGTGTTCAAAAGTTT
GTGACCTTGAGCAACGTGTTAAAGCATTTTGATCAGACGAGGCTACAAAGACAGATTGCAGATATTCATG
TTGCTTTTCAGAGTATGGTAAAGAAAACTGGAGATTGGAAGAAGCATGTGGAAACCAACAGTCGCTTGAT
GAAGAAGTTTGAGGAGTCTCGAGCAGAGTTGGAGAAGGTACTGCGGATTGCTCAGGAGGGCCTGGAGGAA
AAGGGGGATCCAGAGGAGCTCCTGCGGAGACACACTGAGTTTTTCAGTCAGCTGGATCAGAGGGTGCTCA
ATGCTTTCCTGAAAGCTTGTGATGAACTCACCGACATCCTTCCAGAGCAGGAGCAGCAGGGGCTGCAGGA
AGCTGTTCGAAAGCTCCACAAACAATGGAAGGATCTTCAAGGAGAAGCCCCTTATCATTTGCTTCATCTG
AAGATTGATGTGGAGAAGAATAGGTTCTTAGCCTCTGTAGAAGAATGCAGAACTGAGCTGGATCGAGAGA
CCAAGCTGATGCCCCAGGAAGGCAGTGAAAAGATAATTAAAGAGCACAGGGTTTTCTTCAGTGACAAAGG
TCCTCATCATCTCTGTGAGAAAAGGTTACAGCTCATCGAGGAACTCTGTGTGAAACTCCCAGTGCGGGAC
CCAGTAAGGGACACACCTGGAACCTGTCACGTGACTCTCAAAGAGCTCAGAGCTGCCATTGACAGCACCT
ACAGGAAGCTCATGGAAGACCCAGACAAGTGGAAGGACTACACTAGCAGATTCTCTGAGTTCTCATCTTG
GATATCTACAAATGAGACACAATTAAAGGGGATCAAGGGTGAGGCCATCGATACTGCCAACCACGGAGAG
GTTAAACGTGCCGTTGAAGAGATCAGAAATGGTGTTACCAAAAGGGGTGAGACCCTCAGCTGGCTGAAAT
CCAGGCTGAAAGTTTTGACAGAAGTTTCTTCTGAGAATGAAGCCCAAAAGCAGGGAGATGAGCTGGCAAA
ATTATCCAGCTCTTTCAAGGCTCTTGTGACGCTGCTGTCAGAGGTTGAAAAGATGCTAAGCAATTTTGGG
GACTGTGTCCAGTACAAAGAAATAGTCAAAAAATTCTCTCGAAGAATTAATTTCTGGCTCTAAAGAAGTCC
AGGAACAAGCTGAGAAGATCTTGGATACTGAAAATCTGTTTGAAGCACAGCAGTTACTTCTTCATCACCA
GCAAAAGACAAAGCGGATCTCAGCAAAGAAGAGAGATGTGCAGCAGCAGATCGCGCAGGCGCAGCAGGGA
GAAGGGGGGCTGCCTGACCGAGGCCACGAGGAGCTGCGGAAGCTGGAGAGCACACTGGATGGCCTGGAGC
GCAGCCGGGAGAGGCAGGAACGCCGCATCCAGGTCACATTAAGAAAATGGGAGCGATTTGAAACAAACAA
AGAAACAGTAGTAAGATACCTTTTTCAAACAGGTTCCAGTCATGAACGCTTCTTGAGTTTTAGCAGTTTG
GAAAGTTTATCTTCAGAACTGGAACAAACAAAGGAGTTTTCTAAACGGACAGAAAGTATTGCAGTCCAGG
CTGAGAACCTTGTAAAGGAAGCTTCAGAGATACCGCTTGGGCCCCAAAATAAGCAGCTGCTTCAACAGCA
GGCCAAGTCAATCAAAGAACAAGTCAAAAAATTAGAAGACACGCTTGAAGAAGATATTAAACCCATGGAA
ATGGTGAAAACCAAGTGGGATCATTTTGGCAGTAATTTTGAGACTCTGTCCGTCTGGATAACTGAGAAAG
AAAAAGAACTCAATGCCTTGGAAACTTCGTCATCTGCCATGGACATGCAAATCAGCCAAATTAAGGTCAC
AATTCAGGAAATAGAAAGTAAGCTCAGCAGCATTGTAGGATTAGAAGAAGAAGCCCAGTCTTTTGCTCAG
TTTGTTACCACTGGAGAATCTGCTCGAATTAAAGCCAAGTTGACACAAATAAGAAGATACGGGGAAGAGC
TTCGAGAGCATGCACAGTGTCTGGAAGGAACAATCCTGGGACATTTATCTCAGCAGCAAAAGTTTGAAGA
GAACCTTAGAAAGATCCAGCAATCTGTGTCTGAATTTGAAGATAAACTTGCTGTTCCAATTAAAATATGT
TCTTCAGCTACAGAAACATACAAAGTTCTTCAAGAACATATGGATCTCTGCCAGGCCCTGGAGTCACTGA
GCAGCGCGATCACTGCCTTCTCAGCCAGTGCCAGGAAGGTTGTGAACAGAGATTCCTGTGTTCAGGAGGC
TGCGGCTCTACAGCAGCAATACGAGGACATCCTAAGGAGGGCGAAGGAGAGACAGACGGCGCTGGAGAAT
CTGCTGGCCCACTGGCAGAGGCTAGAGAAAGAACTATCATCCTTTTTGACCTGGTTAGAGCGGGGTGAAG
CTAAAGCCAGTTCCCCAGAAATGGACATTTCTGCAGACAGAGTCAAAGTGGAAGGTGAACTTCAGTTAAT
ACAGGCAAGTTCAAGGAAGTGTGAGGAAGGAAAAAATAAAATGCTTTTTGTTACAGTTACATTATTTAAA
ATAATAAAATAAATAAAACTTGTAAAAAAAAAAAAAAAAAAAAAAAA

Human SYNE1 mRNA sequence - var9 (public gi: 17861377) (SEQ ID NO: 191)
AAGGTAAAGCCACTAGAGAGAAACTGAAAGAAAACATTCTTAAGATAAATTGAATTGACATTTTCTCTCT
AAAATATGATTTATAGACCACAGATAGGAATTAAGAGTTTCCTGATAATTTTGGCTTCATATTATTTTAA
AGGATTATCAAGAGGAAATTGCTATTGCTCAAGAGAACAAAATACAGCTCCAACAAATGGGAGAACGACT
TGCTAAAGCCAGCCATGAAAGCAAAGCATCTGAGATTGAATACAAGCTGGGAAAGGTCAACGACCGGTGG
CAGCATCTCCTGGACCTCATTGCAGCCAGGGTGAAGAAGCTGAAGGAGACCCTGGTAGCCGTGCAGCAGC
TTGATAAGAACATGAGCAGCCTGAGGACCTGGCTCGCTCACATCGAGTCAGAGCTGGCCAAGCCAATAGT

CTACGATTCCTGTAACTCGGAAGAAATACAGAGAAAGCTTAATGAGCAGCAGGAGCTTCAGAGAGACATA
GAGAAGCACAGTACAGGTGTTGCATCTGTCCTCAACCTGTGTGAAGTCCTGCTGCACGACTGTGACGCCT
GTGCCACTGATGCCGAGTGTGACTCTATACAGCAGGCTACGAGAAACCTGGACCGGCGGTGGAGAAACAT
TTGTGCTATGTCCATGGAAAGGAGGCTGAAAATCGAAGAGACGTGGCGATTGTGGCAGAAATTTCTGGAT
GACTATTCACGTTTTGAAGATTGGCTGAAGTCTTCAGAAAGGACAGCTGCTTTTCCCAGCTCTTCTGGGG
TGATCTATACAGTTGCCAAGGAAGAACTAAAGAAATTTGAGGCTTTCCAGCGACAGGTCCACGAGTGCCT
GACGCAGCTGGAACTGATCAACAAGCAGTACCGCCGCCTGGCCAGGGAGAACCGCACTGATTCAGCATGT
AGCCTCAAACAGATGGTTCACGAAGGCAACCAGAGATGGGACAACCTGCAAAAGCGTGTCACCTCCATCT
TGCGCAGACTCAAGCATTTTATTGGCCAGCGTGAGGAGTTTGAGACTGCGCGGGACAGCATTCTGGTCTG
GCTCACAGAGATGGATCTGCAGCTCACTAATATTGAACATTTTTCTGAGTGTGATGTTCAAGCTAAAATA
AAGCAACTCAAGGCCTTCCAGCAGGAAATTTCACTGAACCACAATAAGATTGAGCAGATAATTGCCCAAG
GAGAACAGCTGATAGAAAAGAGTGAGCCCTTGGATGCAGCGATCATCGAGGAGGAACTAGATGAGCTCCG
ACGGTACTGCCAGGAGGTCTTCGGGCGTGTGGAAAGATACCATAAGAAACTGATCCGCCTGCCTCTCCCA
GACGATGAGCACGACCTCTCAGACAGGGAGCTGGAGCTGGAAGACTCTGCAGCTCTGTCGGACCTGCACT
GGCACGACCGCTCTGCAGACAGCCTGCTTTCTCCACAGCCTTCCTCCAATCTCTCCCTCTCGCTCGCTCA
GCCCCTCCGGAGCGAGCGGTCAGGACGAGACACCCCAGCTAGTGTGGACTCCATCCCCCTGGAGTGGGAT
CACGACTATGACCTCAGTCGGGACCTGGAGTCTGCAATGTCCAGAGCTCTGCCCTCTGAGGATGAAGAAG
GTCAGGATGACAAAGATTTCTACCTCCGGGGAGCTGTTGCCTTATCAGATGTAATGATCCCCGAAAGCCC
TGAGGCCTATGTAAAACTCACAGAAAATGCAATCAAAAATACCTCCGGGGACCACAGTGCCCTAGAGTCA
CAGATCCGACAACTGGGCAAAGCCCTGGATGATAGCCGCTTTCAGATACAGCAAACCGAAAATATCATTC
GCAGCAAAACTCCCACGGGGCCGGAGCTAGACACCAGCTACAAAGGCTACATGAAACTGCTGGGCGAATG
CAGTAGCAGTATAGACTCCGTGAAGAGACTGGAGCACAAACTGAAGGAGGAAGAGGAGAGCCTTCCTGGC
TTTGTTAACCTGCATAGTACCGAAACCCAAACGGCTGGTGTGATTGACCGATGGGAGCTTCTCCAGGCCC
AGGCATTGAGCAAGGAGTTGAGGATGAAGCAGAACCTCCAGAAGTGGCAGCAGTTTAACTCAGACTTGAA
CAGCATCTGGGCCTGGCTGGGGGACACGGAGGAGGAGTTGGAACAGCTCCAGCGTCTGGAACTCAGCACT
GACATCCAGACCATCGAGCTCCAGATCAAAAAGCTCAAGGAGCTCCAGAAAGCTGTGGACCACCGCAAAG
CCATCATCCTCTCCATCAATCTCTGCAGCCCTGAGTTCACCCAGGCTGACAGCAAGGAGAGCCGGGACCT
GCAGGATCGCTTGTCGCAGATGAATGGGCGCTGGGACCGAGTGTGCTCTCTGCTGGAGGAGTGGCGGGGC
CTGCTGCAGGATGCCCTGATGCAGTGCCAGGGTTTCCATGAAATGAGCCATGGTTTGCTTCTTATGCTGG
AGAACATTGACAGAAGGAAAAATGAAATTGTCCCTATTGATTCTAACCTTGATGCAGAGATACTTCAGGA
CCATCACAAACAGCTTATGCAAATAAAGCATGAGCTGTTGGAATCCCAACTCAGAGTAGCCTCTTTGCAA
GACATGTCTTGCCAACTACTGGTGAATGCTGAAGGAACAGACTGTTTAGAAGCCAAAGAAAAAGTCCATG
TTATTGGAAATCGGCTCAAACTTCTCTTGAAGGAGGTCAGTCGTCATATCAAGGAACTGGAGAAGTTATT
AGACGTGTCAAGTAGTCAGCAGGATTTGTCTTCCTGGTCTTCTGCTGATGAACTGGACACCTCAGGGTCT
GTGAGTCCCACATCAGGAAGGAGCACCCCAAACAGACAGAAAACGCCACGAGGCAAGTGTAGTCTCTCAC
AGCCTGGACCCTCTGTCAGCAGTCCACATAGCAGGTCCACAAAAGGTGGCTCCGATTCCTCCCTTTCTGA
GCCAGGGCCAGGTCGGTCCGGCCGCGGCTTCCTGTTCAGAGTCCTCCGAGCAGCTCTTCCCCTTCAGCTT
CTCCTGCTCCTCCTCATCGGGCTTGCCTGCCTTGTACCAATGTCAGAGGAAGACTACAGCTGTGCCCTCT
CCAACAACTTTGCCCGGTCATTCCACCCCATGCTCAGATACACGAATGGCCCTCCTCCACTCTGAACTAA
GCAGATGCCATCTGCAGAAGTGCTGGTAGCATAAGGAGGATCGGGTCATAAGCAATCCCAAACTACCAAC
AAGAGGACCTTGATCTTGGCGAAAGCCCTCGGTGTGGCAGCTTTAGCCCTCCTCCAGATCACATGTGTGC
AAATTATGGCTTCAGAGGTGGAAGATAAACAGTGACGGGGGAACAAACAGACAACAAGAAGGTTTGGAAG
AAATCTGGTTTGAGACTCTGAACCTTAGCACTAAGGAGATTGAGTAAGGACCTCCAAAGTTCCCCGGACT
CATGAATTCTGGGCCCTTGGCCCATTCTGTGCACAGCCAAGGACTTCAGTAGACCATCTGGGCAGCTTTC
CCATGGTGCTGCTCCAACCATCAGATAAATGACCCTCCCAAGCACCATGTCAGTGTCGTACAATCTACCA
ACCAACCAGTGCTGAAGAGATTTTAGAACCTTGTAACATACAATTTTTAAGAGCTTATATGGCAGCTTCC
TTTTTACCTTGTTTTCCTTTGGGGCATGATGTTTTAACCTTTGCTTTAGAAGCACAAGCTGTAAATCTAA
AAGGCACTTTTTTTAGAGGTATAAAGAAAAACTAGATGTAATAAATAAGATCATGGAAGGCTTTATGTG
AAAAAAGTTGAATGTTATAGT

Human SYNE1 mRNA sequence - var10 (public gi: 17861385) (SEQ ID NO: 192)

CAAAAATCAGTCTGATCTCGGGAAACCTGGAGAAATTTATTTCTGTACTCTAATGTTCTTTCATTTTGG
TGACCATCAAGGTGCTGGGAGAGGAATTAGATGGCTGTAATTCAAAGTTAATGGAATTAGATGCAGCAGT
ACAGAAATTCTTGGAACAGAATGGCCAACTGGGTAAGCCACTGGCCAAGAAGATAGGAAAACTGACTGAA
CTTCACCAGCAGACCATTAGACAAGCTGAGAATCGGCTCTCCAAGCTCAATCAGGCAACATCACATTTAG
AAGAATACAATGAAATGCTTGAATTAATTTTGAAGTGGATTGAAAAAGCTAAAGTCTTGGCTCATGGAAC
TATTGCATGGAATTCTGCAAGCCAGCTTCGGAAACAATATATTTGCATCAGACCCTGCTAGAAGAATCC
AAAGAAATTGACAGTGAGCTGGAAGCAATGACTGAGAAATTACAGTACCTCACTAGCGTGTACTGTACAG
AAAAAATGTCTCAGCAAGTGGCAGAACTGGGACGGGAGACTGAGGAGTTGCGACAGATGATCAAAATTCG
TTTGCAGAACCTCCAAGATGCAGCTAAGGATATGAAAAAATTTGAAGCAGAGTTGAAAAAGTTACAAGCT
GCCTTGGAGCAAGCCCAGGCAACACTGACTTCTCCAGAAGTTGGACGTCTCAGTCTCAAGGAGCAGCTCT
CTCATCGGCAGCATTTGTTGTCTGAGATGGAGTCACTGAAGCCGAAGGTGCAAGCAGTGCAGCTCTGCCA
GAGTGCCCTCCGGATCCCCGAGGATGTGGTTGCCAGCTTACCTCTCTGTCATGCTGCTCTGCGGCTGCAG

GAAGAGGCCAGCCGGCTGCAGCACACCGCCATCCAGCAGTGTAACATCATGCAGGAAGCTGTGGTACAAT
ATGAACAATATGAGCAAGAAATGAAACATCTCCAGCAACTGATAGAAGGAGCTCACAGAGAGATTGAGGA
TAAACCTGTTGCCACCAGTAACATACAGGAGCTGCAGGCTCAGATTTCTCGGCATGAGGAGCTGGCGCAG
AAAATTAAGGGCTACCAGGAGCAGATCGCTTCTTTGAATTCCAAGTGCAAGATGCTGACGATGAAAGCCA
AGCACGCCACCATGCTGCTGACGGTGACCGAGGTCGAGGGGCTGGCGGAAGGGACAGAGGACCTGGATGG
GGAGCTCCTCCCCACGCCTTCGGCCCACCCCTCTGTGGTCATGATGACTGCAGGTCGCTGTCACACTTTG
CTGTCACCGGTCACTGAGGAGTCTGGGGAGGAGGGAACCAACAGTGAGATTTCCTCTCCACCTGCCTGTC
GCTCCCCTTCACCTGTGGCTAATACAGATGCTTCTGTTAACCAGGACATTGCATATTACCAAGCCTTGTC
TGCTGAGAGGTTGCAGACAGATGCTGCAAAAATTCACCCCAGCACATCCGCATCCCAGGAGTTCTATGAA
CCGGGATTGGAGCCATCCGCTACTGCCAAACTGGGTGATTTGCAGCGTTCTTGGGAAACCTTAAAGAATG
TGATCAGTGAGAAGCAGCGCACACTCTATGAAGCTTTGGAGCGCCAGCAGAAGTACCAGGACTCCCTCCA
GTCCATCTCTACGAAGATGGAGGCCATTGAGCTGAAACTCAGTGAGAGCCCAGAGCCTGGCAGGAGTCCA
GAAAGCCAGATGGCTGAACATCAGGCATTGATGGATGAGATTCTCATGCTCCAGGATGAAATCAATGAGC
TCCAGTCCTCTCTCGCAGAGGAGCTGGTATCCGAGTCTTGTGAGGCCGACCCTGCGGAGCAGCTGGCCTT
GCAGTCCACGCTCACTGTCTTAGCCGAGCGAATGTCCACCATCAGGATGAAAGCCTCGGGGAAACGGCAG
CTTTTGGAGGAGAAGTTGAATGATCAGCTGGAGGAACAAAGGCAGGAACAGGCCCTGCAGAGGTATCGCT
GTGAAGCCGATGAGCTGGACAGCTGGCTCTTGAGTACCAAGGCCACTCTGGACACTGCGCTGAGTCCACC
CAAGGAGCCCATGGACATGGAGGCCCAGCTTATGGACTGCCAGAATATGCTGGTGGAAATAGAGCAGAAG
GTGGTGGCTTTATCAGAACTGTCAGTCCACAATGAGAACCTGCTGCTGGAGGGCAAAGCTCACACCAAGG
ACGAGGCCGAGCAGCTGGCTGGAAAGCTGAGAAGGCTCAAGGGGAGCCTGCTGGAGCTGCAGAGAGCCCT
GCATGATAAGCAGCTCAACATGCAGGGAACAGCACAGGAGAAGGAGGAGAGCGATGTTGACCTAACAGCC
ACGCAGAGCCCCGGCGTCCAGGAATGGCTGGCCCAAGCTCGCACCACATGGACCCAGCAGCGGCAGAGCA
GTCTCCAGCAACAAAAAGAGTTAGAACAGGAATTAGCCGAGCAGAAGAGTCTCCTTCGCTCAGTAGCCAG
TCGTGGAGAGGAGATTCTAATTCAACATTCGGCGGCAGAGACCTCTGGTGATGCTGGCGAAAAACCTGAT
GTGTTATCCCAGGAGTTGGGGATGGAAGGGGAGAAATCATCCGCTGAAGACCAGATGAGAATGAAATGGG
AAAGCCTACATCAAGAATTTAGTACCAAGCAGAAACTACTACAGAATGTTCTGGAACAGGAACAAGAGCA
AGTGCTTTATAGCAGGCCAAATCGACTCTTGTCTGGTGTGCCACTGTACAAAGGGGACGTGCCAACCCAA
GATAAATCTGCAGTTACATCTTTGCTGGATGGACTGAACCAAGCCTTCGAGGAGGTTTCATCCCAGAGTG
GAGGGGCAAAGAGGCAGAGTATACACTTGGAGCAGAAGTTGTATGATGGAGTCTCAGCCACCTCTACTTG
GTTGGATGACGTTGAAGAACGTTTATTTGTTGCCACAGCACTTTTACCAGAAGAAACAGAGACTTGTCTC
TTCAACCAAGAGATTCTTGCCAAAGACATTAAGGAAATGTCTGAAGAAATGGATAAGAACAAAAACTTGT
TTTCCCAAGCTTTTCCAGAGAATGGTGATAATCGAGATGTTATTGAAGATACTTTGGGTTGTCTTTTGGG
CAGGTTATCCTTGCTAGACTCAGTAGTGAATCAACGATGTCATCAGATGAAAGAAAGACTTCAGCAAATA
CTAAATTTCCAGAATGATCTGAAAGTGCTGTTTACATCACTGGCTGACAACAAATACATCATTCTGCAAA
AACTGGCAAATGTGTTTGAACAGCCCGTAGCAGAACAAATAGAGGCAATACAACAGGCTGAAGATGGACT
CAAAGAATTTGATGCAGGAATCATTGAATTAAAGAGGCGTGGTGACGAGCTACAGGTCGAGCAGCCGTCC
ATGCAAGAACTCTCCAAGCTCCAGGACATGTATGATGAGCTGATGATGATCATTGGCTCCCGGAGGAGTG
GTCTGAATCAGAACCTTACACTCAAGAGTCAGTATGAGAGGGCCCTACAAGATCTGGCTGACCTGCTAGA
AACTGGTCAGGAGAAGATGGCAGGAGACCAGAAAATCATCGTGTCTTCCAAAGAGGAAATCCAGCAACCA
CTTGACAAACATAAGGAATACTTTCAGGGCCTGGAATCTCATATGATCTTGACTGTAACACTCTTCAGAA
AGATAATCAGCTTTGCAGTCCAAAAGGAAACCCAGTTCCATACAGAGCTGATGGCTCAGGCTTCTGCTGT
ACTGAAACGGGCTCACAAGAGGGGTGTGGAGCTGGAGTACATTCTAGAGACGTGGTCCCATCTGGATGAG
GACCAGCAGGAGCTCAGCAGACAGCTGGAGGTGGTGGAAAGCAGCATCCCAAGCGTGGGTCTGGTGGAGG
AGAACGAGGACAGGCTTATTGACCGCATAACACTCTACCAGCATTTAAAATCTAGCCTTAATGAATACCA
GCCCAAATTATATCAAGTATTAGATGATGGGAAACGACTTCTGATATCCATCAGCTGCTCAGATCTAGAA
AGCCAACTAAATCAACTTGGAGAGTGCTGGCTAAGTAACACCAATAAAATGTCTAAGGAACTTCACAGAC
TGGAAACAATATTGAAACACTGGACCAGATATCAAAGTGAATCTGCAGATCTAATTCACTGGTTACAATC
TGCAAAAGACCGGCTAGAATTTTGGACTCAGCAATCTGTGACAGTCCCACAAGAGCTGGAAATGGTCCGT
GATCATCTAAATGCTTTCCTGGAGTTTTCTAAAGAAGTGGATGCCCAATCTTCCCTGAAATCATCTGTTC
TGAGTACTGGAAATCAGCTCCTTCGACTAAAAAAGGTGGACACAGCCACGCTGCGCTCTGAGTTGTCGCG
CATTGATAGCCAGTGGACTGACCTGCTAACCAATATCCCAGCCGTCCAGGAGAAGCTCCACCAGCTCCAG
ATGGATAAACTGCCTTCCCGCCATGCCATTTCTGAAGTCATGAGTTGGACTTCTCTAATGGAAAATGCTA
TTCAGAAGGATGAAGATAATATTAAAAATTCCATAGGTTACAAGGCAATTCATGAATACCTTCAGAAATA
TAAGGGTTTTAAGATAGACATTAACTGTAAACAGCTGACAGTGGATTTTGTGAACCAGTCCGTGCTACAA
ATCAGCAGTCAGGATGTGGAAAGTAAGCGTAGTGATAAGACTGATTTTGCTGAGCAACTTGGAGCAATGA
ATAAAAGTTGGCAAATTCTGCAAGGTCTAGTAACTGAGAAGATCCAGCTGTTGGAAGGCTTATTGGAATC
TTGGTCAGAATATGAAAATAATGTACAATGTCTGAAAACATGGTTTGAAACCCAGGAAAAGAGACTAAAA
CAACAGCATCGAATTGGAGATCAGGCTTCTGTTCAAAATGCACTGAAAGACTGTCAGGATCTGGAAGATC
TGATTAAAGCAAAAGATAAAGAAGTAGAGAAAATTGAGCAGAATGGACTTGCTTTGATTCAGACCAAGAA
AGAAGACGTCTCTAGCATTGTCATGAGCACACTGCGAGAGCTCGGCCAAACCTGGGCAAATTTAGATCAC
ATGGTTGGACAATTAAAGATACTGCTGAAATCAGTGCTTGACCAATGGAGTAGTCACAAAGTGGCCTTTG
ACAAGATAAACAGTTACCTCATGGAGGCCAGATACTCTCTTTCCCGATTCCGTCTGCTGACTGGCTCCTT
AGAAGCTGTGCAAGTTCAGGTGGACAATCTTCAGAATCTCCAAGATGATCTGGAAAAACAGGAAAGGAGC

Figure 36 part - 105

TTACAGAAATTTGGCTCTATCACCAACCAATTATTAAAAGAGTGTCACCCACCCGTGACAGAAACTCTTA
CCAATACACTGAAAGAAGTCAACATGAGATGGAATAACTTGCTGGAAGAGATTGCTGAGCAGCTACAGTC
CAGCAAGGCCCTACTTCAGCTTTGGCAAAGATACAAGGACTACTCCAAACAGTGTGCTTCGACAGTTCAG
CAGCAGGAGGATCGAACCAATGAGCTGTTGAAGGCAGCCACAAACAAGGACATTGCCGATGATGAGGTTG
CCACATGGATTCAAGATTGCAACGACCTCCTCAAAGGACTGGGCACAGTTAAAGATTCCCTCTTTGTTCT
CCATGAGCTGGGAGAGCAACTGAAGCAACAAGTGGATGCTTCCGCAGCATCAGCTATTCAATCGGATCAA
CTCTCTTTGAGTCAACACTTGTGTGCCCTGGAGCAAGCTCTCTGCAAACAGCAGACTTCATTACAGGCTG
GAGTTCTTGATTATGAAACCTTTGCCAAGAGTTTAGAAGCTTTGGAGGCCTGGATAGTGGAAGCTGAAGA
AATACTACAAGGGCAGGACCCTAGCCACTCATCTGACCTCTCCACAATCCAGGAAAGGATGGAAGAACTT
AAGGGACAGATGTTAAAATTCAGCAGCATGGCTCCAGATTTAGACCGTCTAAATGAGCTTGGATATAGGT
TACCCTTGAATGATAAGGAAATCAAAAGAATGCAGAATCTGAACCGCCATTGGTCTCTGATCTCCTCTCA
GACTACAGAAAGATTCAGCAAGTTGCAGTCATTTTTGCTACAACATCAGACTTTCTTGGAAAAATGTGAA
ACATGGATGGAATTCCTAGTTCAGACAGAACAAAAGTTAGCAGTAGAGATTTCAGGAAATTATCAGCACC
TTTTGGAACAGCAGAGAGCACACGAGTTGTTTCAAGCCGAGATGTTCAGTCGTCAGCAGATTTTGCACTC
AATCATTATTGATGGGCAACGTCTTCTAGAACAAGGTCAAGTTGATGACAGGGATGAATTCAACCTGAAA
TTGACACTCCTCAGTAATCAATGGCAGGGAGTGATTCGCAGGGCCCAGCAGAGGCGGGGGATCATTGACA
GCCAGATTCGCCAGTGGCAGCGCTATAGGGAGATGGCAGAAAAGCTTCGTAAATGGTTGGTTGAAGTGTC
CTACCTCCCCATGAGTGGTCTCGGAAGTGTTCCTATACCACTGCAACAAGCAAGGACCCTCTTTGATGAA
GTGCAGTTCAAAGAAAAAGTGTTTCTGCGGCAACAAGGCAGCTACATCCTGACTGTGGAGGCTGGCAAGC
AACTCCTTCTCTCGGCGGACAGTGGCGCTGAGGCCGCCTTGCAGGCCGAACTCGCTGAAATCCAAGAGAA
ATGGAAATCAGCCAGCATGCGGCTGGAAGAACAGAAGAAAAAACTAGCCTTCTTGTTGAAAGACTGGGAA
AAATGTGAGAAAGGAATAGCAGATTCCCTGGAGAAACTACGAACTTTCAAAAAGAAGCTTTCGCAGTCTC
TCCCGGATCACCATGAAGAGCTCCATGCAGAACAAATGCGTTGCAAGGAATTAGAAAATGCAGTTGGGAG
CTGGACAGATGACTTGACCCAGTTGAGCCTGCTGAAGGACACCCTCTCTGCCTATATCAGTGCTGATGAT
ATCTCCATTCTTAATGAACGCGTAGAGCTTCTGCAAAGGCAGTGGGAAGAACTATGCCACCAGCTCTCCT
TAAGGCGGCAGCAAATAGGTGAAAGATTGAATGAATGGGCAGTCTTCAGTGAAAAGAACAAGGAACTCTG
TGAGTGGTTGACTCAAATGGAAAGCAAAGTTTCTCAGAATGGAGACATTCTCATTGAAGAAATGATAGAG
AAGCTCAAGAAGGATTATCAAGAGGAAATTGCTATTGCTCAAGAGAACAAAATACAGCTCCAACAAATGG
GAGAACGACTTGCTAAAGCCAGCCATGAAAGCAAAGCATCTGAGATTGAATACAAGCTGGGAAAGGTCAA
CGACCGGTGGCAGCATCTCCTGGACCTCATTGCAGCCAGGGTGAAGAAGCTGAAGGAGACCCTGGTAGCC
GTGCAGCAGCTTGATAAGAACATGAGCAGCCTGAGGACCTGGCTCGCTCACATCGAGTCAGAGCTGGCCA
AGCCAATAGTCTACGATTCCTGTAACTCGGAAGAAATACAGAGAAAGCTTAATGAGCAGCAGGAGCTTCA
GAGAGACATAGAGAAGCACAGTACAGGTGTTGCATCTGTCCTCAACCTGTGTGAAGTCCTGCTGCACGAC
TGTGACGCCTGTGCCACTGATGCCGAGTGTGACTCTATACAGCAGGCTACGAGAAACCTGGACCGGCGGT
GGAGAAACATTTGTGCTATGTCCATGGAAAGGAGGCTGAAAATCGAAGAGACGTGGCGATTGTGGCAGAA
ATTTCTGGATGACTATTCACGTTTTGAAGATTGGCTGAAGTCTTCAGAAAGGACAGCTGCTTTTCCCAGC
TCTTCTGGGGTGATCTATACAGTTGCCAAGGAAGAACTAAAGAAATTTGAGGCTTTCCAGCGACAGGTCC
ACGAGTGCCTGACGCAGCTGGAACTGATCAACAAGCAGTACCGCCGCCTGGCCAGGGAGAACCGCACTGA
TTCAGCATGTAGCCTCAAACAGATGGTTCACGAAGGCAACCAGAGATGGGACAACCTGCAAAAGCGTGTC
ACCTCCATCTTGCGCAGACTCAAGCATTTTATTGGCCAGCGTGAGGAGTTTGAGACTGCGCGGGACAGCA
TTCTGGTCTGGCTCACAGAGATGGATCTGCAGCTCACTAATATTGAACATTTTTCTGAGTGTGATGTTCA
AGCTAAAATAAAGCAACTCAAGGCCTTCCAGCAGGAAATTTCACTGAACCACAATAAGATTGAGCAGATA
ATTGCCCAAGGAGAACAGCTGATAGAAAAGAGTGAGCCCTTGGATGCAGCGATCATCGAGGAGGAACTAG
ATGAGCTCCGACGGTACTGCCAGGAGGTCTTCGGGCGTGTGGAAAGATACCATAAGAAACTGATCCGCCT
GCCTCTCCCAGACGATGAGCACGACCTCTCAGACAGGGAGCTGGAGCTGGAAGACTCTGCAGCTCTGTCG
GACCTGCACTGGCACGACCGCTCTGCAGACAGCCTGCTTTCTCCACAGCCTTCCTCCAATCTCTCCCTCT
CGCTCGCTCAGCCCCTCCGGAGCGAGCGGTCAGGACGAGACACCCCAGCTAGTGTGGACTCCATCCCCCT
GGAGTGGGATCACGACTATGACCTCAGTCGGGACCTGGAGTCTGCAATGTCCAGAGCTCTGCCCTCTGAG
GATGAAGAAGGTCAGGATGACAAAGATTTCTACCTCCGGGGAGCTGTTGCCTTATCAGGGGACCACAGTG
CCCTAGAGTCACAGATCCGACAACTGGGCAAAGCCCTGGATGATAGCCGCTTTCAGATACAGCAAACCGA
AAATATCATTCGCAGCAAAACTCCCACGGGGCCGGAGCTAGACACCAGCTACAAAGGCTACATGAAACTG
CTGGGCGAATGCAGTAGCAGTATAGACTCCGTGAAGAGACTGGAGCACAAACTGAAGGAGGAAGAGGAGA
GCCTTCCTGGCTTTGTTAACCTGCATAGTACCGAAACCCAAACGGCTGGTGTGATTGACCGATGGGAGCT
TCTCCAGGCCCAGGCATTGAGCAAGGAGTTGAGGATGAAGCAGAACCTCCAGAAGTGGCAGCAGTTTAAC
TCAGACTTGAACAGCATCTGGGCCTGGCTGGGGGACACGGAGGAGGAGTTGGAACAGCTCCAGCGTCTGG
AACTCAGCACTGACATCCAGACCATCGAGCTCCAGATCAAAAAGCTCAAGGAGCTCCAGAAAGCTGTGGA
CCACCGCAAAGCCATCATCCTCTCCATCAATCTCTGCAGCCCTGAGTTCACCCAGGCTGACAGCAAGGAG
AGCCGGGACCTGCAGGATCGCTTGTCGCAGATGAATGGGCGCTGGGACCGAGTGTGCTCTCTGCTGGAGG
AGTGGCGGGGCCTGCTGCAGGATGCCCTGATGCAGTGCCAGGGTTTCCATGAAATGAGCCATGGTTTGCT
TCTTATGCTGGAGAACATTGACAGAAGGAAAAATGAAATTGTCCCTATTGATTCTAACCTTGATGCAGAG
ATACTTCAGGACCATCACAAACAGCTTATGCAAATAAAGCATGAGCTGTTGGAATCCCAACTCAGAGTAG
CCTCTTTGCAAGACATGTCTTGCCAACTACTGGTGAATGCTGAAGGAACAGACTGTTTAGAAGCCAAAGA
AAAAGTCCATGTTATTGGAAATCGGCTCAAACTTCTCTTGAAGGAGGTCAGTCGTCATATCAAGGAACTG

GAGAAGTTATTAGACGTGTCAAGTAGTCAGCAGGATTTGTCTTCCTGGTCTTCTGCTGATGAACTGGACA
CCTCAGGGTCTGTGAGTCCCACATCAGGAAGGAGCACCCCAAACAGACAGAAAACGCCACGAGGCAAGTG
TAGTCTCTCACAGCCTGGACCCTCTGTCAGCAGTCCACATAGCAGGTCCACAAAAGGTGGCTCCGATTCC
TCCCTTTCTGAGCCAGGGCCAGGTCGGTCCGGCCGCGGCTTCCTGTTCAGAGTCCTCCGAGCAGCTCTTC
CCCTTCAGCTTCTCCTGCTCCTCCTCATCGGGCTTGCCTGCCTTGTACCAATGTCAGAGGAAGACTACAG
CTGTGCCCTCTCCAACAACTTTGCCCGGTCATTCCACCCCATGCTCAGATACACGAATGGCCCTCCTCCA
CTCTGAACTAAGCAGATGCCATCTGCAGAAGTGCTGGTAGCATAAGGAGGATCGGGTCATAAGCAATCCC
AAACTACCAACAAGAGGACCTTGATCTTGGCGAAAGCCCTCGGTGTGGCAGCTTTAGCCCTCCTCCAGAT
CACATGTGTGCAAATTATGGCTTCAGAGGTGGAAGATAAACAGTGACGGGGAACAAACAGACAACAAGA
AGGTTTGGAAGAAATCTGGTTTGAGACTCTGAACCTTAGCACTAAGGAGATTGAGTAAGGACCTCCAAAG
TTCCCCGGACTCATGAATTCTGGGCCCTTGGCCCATTCTGTGCACAGCCAAGGACTTCAGTAGACCATCT
GGGCAGCTTTCCCATGGTGCTGCTCCAACCATCAGATAAATGACCCTCCCAAGCACCATGTCAGTGTCGT
ACAATCTACCAACCAACCAGTGCTGAAGAGATTTTAGAACCTTGTAACATACAATTTTTAAGAGCTTATA
TGGCAGCTTCCTTTTTACCTTGTTTTCCTTTGGGGCATGATGTTTTAACCTTTGCTTTAGAAGCACAAGC
TGTAAATCTAAAAGGCACTTTTTTTTAGAGGTATAAAGAAAAACTAGATGTAATAAATAAGATCATGGAA
GGCTTTATGTGAAAAAAGTTGAATGTTATAGT

Human SYNE1 mRNA sequence - var11 (public gi: 17227153) (SEQ ID NO: 193)
AACTCCTTCTCTCGGCGGACAGTGGCGCTGAGGCCGCCTTGCAGGCCGAACTCGCTGAAATCCAAGAGAA
ATGGAAATCAGCCAGCATGCGGCTGGAAGAACAGAAGAAAAAACTAGCCTTCTTGTTGAAAGACTGGGAA
AAATGTGAGAAAGGAATAGCAGATTCCCTGGAGAAACTACGAACTTTCAAAAAGAAGCTTTCGCAGTCTC
TCCCGGATCACCATGAAGAGCTCCATGCAGAACAAATGCGTTGCAAGGAATTAGAAAATGCAGTTGGGAG
CTGGACAGATGACTTGACCCAGTTGAGCCTGCTGAAGGACACCCTCTCTGCCTATATCAGTGCTGATGAT
ATCTCCATTCTTAATGAACGCGTAGAGCTTCTGCAAAGGCAGTGGAAGAACTATGCCACCAGCTCTCCT
TAAGGCGGCAGCAAATAGGTGAAAGATTGAATGAATGGGCAGTCTTCAGTGAAAAGAACAAGGAACTCTG
TGAGTGGTTGACTCAAATGGAAAGCAAAGTTTCTCAGAATGGAGACATTCTCATTGAAGAAATGATAGAG
AAGCTCAAGAAGGATTATCAAGAGGAAATTGCTATTGCTCAAGAGAACAAAATACAGCTCCAACAAATGG
GAGAACGACTTGCTAAAGCCAGCCATGAAAGCAAAGCATCTGAGATTGAATACAAGCTGGGAAAGGTCAA
CGACCGGTGGCAGCATCTCCTGGACCTCATTGCAGCCAGGGTGAAGAAGCTGAAGGAGACCCTGGTAGCC
GTGCAGCAGCTTGATAAGAACATGAGCAGCCTGAGGACCTGGCTCGCTCACATCGAGTCAGAGCTGGCCA
AGCCAATAGTCTACGATTCCTGTAACTCGGAAGAAATACAGAGAAAGCTTAATGAGCAGCAGGAGCTTCA
GAGAGACATAGAGAAGCACAGTACAGGTGTTGCATCTGTCCTCAACCTGTGTGAAGTCCTGCTGCACGAC
TGTGACGCCTGTGCCACTGATGCCGAGTGTGACTCTATACAGCAGGCTACGAGAAACCTGGACCGGCGGT
GGAGAAACATTTGTGCTATGTCCATGGAAAGGAGGCTGAAAATCGAAGAGACGTGGCGATTGTGGCAGAA
ATTTCTGGATGACTATTCACGTTTTGAAGATTGGCTGAAGTCTTCAGAAAGGACAGCTGCTTTTCCCAGC
TCTTCTGGGGTGATCTATACAGTTGCCAAGGAAGAACTAAAGAAATTTGAGGCTTTCCAGCGACAGGTCC
ACGAGTGCCTGACGCAGCTGGAACTGATCAACAAGCAGTACCGCCGCCTGGCCAGGGAGAACCGCACTGA
TTCAGCATGTAGCCTCAAACAGATGGTTCACGAAGGCAACCAGAGATGGGACAACCTGCAAAAGCGTGTC
ACCTCCATCTTGCGCAGACTCAAGCATTTTATTGGCCAGCGTGAGGAGTTTGAGACTGCGCGGGACAGCA
TTCTGGTCTGGCTCACAGAGATGGATCTGCAGCTCACTAATATTGAACATTTTTCTGAGTGTGATGTTCA
AGCTAAAATAAAGCAACTCAAGGCCTTCCAGCAGGAAATTTCACTGAACCACAATAAGATTGAGCAGATA
ATTGCCCAAGGAGAACAGCTGATAGAAAAGAGTGAGCCCTTGGATGCAGCGATCATCGAGGAGGAACTAG
ATGAGCTCCGACGGTACTGCCAGGAGGTCTTCGGGCGTGTGGAAAGATACCATAAGAAACTGATCCGCCT
GCCTCTCCCAGACGATGAGCACGACCTCTCAGACAGGGAGCTGGAGCTGGAAGACTCTGCAGCTCTGTCG
GACCTGCACTGGCACGACCGCTCTGCAGACAGCCTGCTTTCTCCACAGCCTTCCTCCAATCTCTCCCTCT
CGCTCGCTCAGCCCCTCCGGAGCGAGCGGTCAGGACGAGACACCCCAGCTAGTGTGGACTCCATCCCCCT
GGAGTGGGATCACGACTATGACCTCAGTCGGGACCTGGAGTCTGCAATGTCCAGAGCTCTGCCCTCTGAG
GATGAAGAAGGTCAGGATGACAAAGATTTCTACCTCCGGGGAGCTGTTGCCTTATCAGGGGACCACAGTG
CCCTAGAGTCACAGATCCGACAACTGGGCAAAGCCCTGGATGATAGCCGCTTTCAGATACAGCAAACCGA
AAATATCATTCGCAGCAAAACTCCCACGGGGCCGGAGCTAGACACCAGCTACAAAGGCTACATGAAACTG
CTGGGCGAATGCAGTAGCAGTATAGACTCCGTGAAGAGACTGGAGCACAAACTGAAGGAGGAAGAGGAGA
GCCTTCCTGGCTTTGTTAACCTGCATAGTACCGAAACCCAAACGGCTGGTGTGATTGACCGATGGGAGCT
TCTCCAGGCCCAGGCATTGAGCAAGGAGTTGAGGATGAAGCAGAACCTCCAGAAGTGGCAGCAGTTTAAC
TCAGACTTGAACAGCATCTGGGCCTGGCTGGGGGACACGGAGGAGGAGTTGGAACAGCTCCAGCGTCTGG
AACTCAGCACTGACATCCAGACCATCGAGCTCCAGATCAAAAAGCTCAAGGAGCTCCAGAAAGCTGTGGA
CCACCGCAAAGCCATCATCCTCTCCATCAATCTCTGCAGCCCTGAGTTCACCCAGGCTGACAGCAAGGAG
AGCCGGGACCTGCAGGATCGCTTGTCGCAGATGAATGGGCGCTGGGACCGAGTGTGCTCTCTGCTGGAGG
AGTGGCGGGGCCTGCTGCAGGATGCCCTGATGCAGTGCCAGGGTTTCCATGAAATGAGCCATGGTTTGCT
TCTTATGCTGGAGAACATTGACAGAAGGAAAAATGAAATTGTCCCTATTGATTCTAACCTTGATGCAGAG
ATACTTCAGGACCATCACAAACAGCTTATGCAAATAAAGCATGAGCTGTTGGAATCCCAACTCAGAGTAG
CCTCTTTGCAAGACATGTCTTGCCAACTACTGGTGAATGCTGAAGGAACAGACTGTTTAGAAGCCAAAGA
AAAAGTCCATGTTATTGGAAATCGGCTCAAACTTCTCTTGAAGGAGGTCAGTCGTCATATCAAGGAACTG
GAGAAGTTATTAGACGTGTCAAGTAGTCAGCAGGATTTGTCTTCCTGGTCTTCTGCTGATGAACTGGACA

CCTCAGGGTCTGTGAGTCCCACATCAGGAAGGAGCACCCCAAACAGACAGAAAACGCCACGAGGCAAGTG
TAGTCTCTCACAGCCTGGACCCTCTGTCAGCAGTCCACATAGCAGGTCCACAAAAGGTGGCTCCGATTCC
TCCCTTTCTGAGCCAGGGCCAGGTCGGTCCGGCCGCGGCTTCCTGTTCAGAGTCCTCCGAGCAGCTCTTC
CCCTTCAGCTTCTCCTGCTCCTCCTCATCGGGCTTGCCTGCCTTGTACCAATGTCAGAGGAAGACTACAG
CTGTGCCCTCTCCAACAACTTTGCCCGGTCATTCCACCCCATGCTCAGATACACGAATGGCCCTCCTCCA
CTCTGAACTAAGCAGATGCCATCTGCAGAAGTGCTGGTAGCATAAGGAGGATCGGGTCATAAGCAATCCC
AAACTACCAACAAGAGGACCTTGATCTTGGCGAAAGCCCTCGGTGTGGCAGCTTTAGCCCTCCTCCAGAT
CACATGTGTGCAAATTATGGCTTCAGAGGTGGAAGATAAACAGTGACGGGGGAACAAACAGACAACAAGA
AGGTTTGGAAGAAATCTGGTTTGAGACTCTGAACCTTAGCACTAAGGAGATTGAGTAAGGACCTCCAAAG
TTCCCCGGACTCATGAATTCTGGGCCCTTGGCCCATTCTGTGCACAGCCAAGGACTTCAGTAGACCATCT
GGGCAGCTTTCCCATGGTGCTGCTCCAACCATCAGATAAATGACCCTCCCAAGCACCATGTCAGTGTCGT
ACAATCTACCAACCAACCAGTGCTGAAGAGATTTTAGAACCTTGTAACATACAATTTTTAAGAGCTTATA
TGGCAGCTTCCTTTTTACCTTGTTTTCCTTTGGGGCATGATGTTTTAACCTTTGCTTTAGAAGCACAAGC
TGTAAATCTAAAAGGCACTTTTTTTTAGAGGTATAAAGAAAAACTAGATGTAATAAATAAGATCATGGAA
GGCTTTATGTGAAAAAAGTTGAATGTTATAGT

Human SYNE1 mRNA sequence - var12 (public gi: 16550165) (SEQ ID NO: 194)
ACAAAAGAGCATTCCATGAAATCTACCGGACCAGGTCTGTTAACGGGATTCCAGTGCCACCTGATCAATT
AGAGGACATGGCCGAGAGGTTTCATTTTGTTTCCTCCACATCAGAGCTACACCTAATGAAAATGGAATTT
TTAGAATTAAAGTACCGTCTGCTCTCACTGCTGGTTCTTGCAGAGTCAAAGCTGAAGTCTTGGATCATTA
AGTACGGGAGGAGAGAGTCAGTGGAGCAGCTTCTACAAAACTACGTGTCTTTTATAGAAAATAGCAAGTT
CTTTGAACAATATGAGGTGACATACCAGATCTTGAAACAGACAGCTGAGATGTATGTCAAAGCAGATGGT
TCAGTGGAAGAAGCTGAGAATGTGATGAAATTCATGAATGAAACCACCGCTCAGTGGAGGAATCTCTCAG
TAGAAGTGAGGAGTGTGAGGAGCATGCTGGAAGAAGTGATCTCTAACTGGGATCGCTATGGCAATACAGT
GGCTAGTCTGCAAGCCTGGCTAGAGGATGCTGAAAAAATGCTCAATCAATCAGAAAATGCCAAAAAGGAT
TTTTTTCGAAATTTACCTCATTGGATTCAGCAGCATACTGCCATGAACGATGCTGGCAATTTTCTAATTG
AAACCTGTGATGAGATGGTTTCCCGTGACCTGAAGCAGCAATTACTGTTGCTAAATGGGCGGTGGAGGGA
GTTGTTTATGGAAGTCAAGCAATATGCTCAAGCTGATGAGATGGACAGAATGAAGAAGGAATACACAGAC
TGTGTTGTTACCCTGTCTGCTTTTGCGACGGAAGCCCATAAGAAACTTTCTGAACCCTTAGAAGTCTCTT
TTATGAATGTCAAGCTATTAATTCAAGACTTGGAGGATATTGAGCAGAGGGTGCCTGTGATGGATGCCCA
ATACAAGATAATTACAAAGACAGCACACCTCATTACCAAAGAAAGCCCCCAAGAAGAAGGAAAAGAAATG
TTTGCGACCATGTCAAAGCTCAAAGAGCAGCTAACCAAGGTCAAAGAATGTTACTCCCCACTCCTTTATG
AGTCTCAGCAGCTGTTGATTCCGTTGGAGGAATTAGAAAAGCAGATGACGTCCTTTTATGACTCACTTGG
GAAAATCAATGAAATTATCACAGTTCTTGAGCGTGAGGCACAATCGAGTGCCCTTTTTAAACAAAAACAT
CAGGAACTGTTAGCTTGTCAAGAAAACTGTAAGAAAACCTTGACACTTATTGAGAAAGGCAGTCAAAGTG
TTCAAAAGTTTGTGACCTTGAGCAACGTGTTAAAGCATTTTGATCAGACGAGGCTACAAAGACAGATTGC
AGATATTCATGTTGCTTTTCAGAGTATGGTAAAGAAAACTGGAGATTGGAAGAAGCATGTGGAAACCAAC
AGTCGCTTGATGAAGAAGTTTGAGGAGTCTCGAGCAGAGTTGGAGAAGGTACTGCGGATTGCTCAGGAGG
GCCTGGAGGAAAAGGGGGATCCAGAGGAGCTCCTGCGGAGACACACTGAGTTTTTCAGTCAGCTGGATCA
GAGGGTGCTCAATGCTTTCCTGAAAGCTTGTGATGAACTCACCGACATCCTTCCAGAGCAGGAGCAGCAG
GGGCTGCAGGAAGCTGTTCGAAAGCTCCACAAACAATGGAAGGTGAGTCAGGACAGGACGGAGACACCGT
GCATCCTCAATGAAGGGAGAAGCTTGAGCGTGTTAAGGTCCAAATGTAAGAGAAATTTAGAAATTCCTGG
AAAGTCACTGTAACTATTTCGCTCATTTAAAAACTCAAAAAACTGGACTTAAATAAAACCTGATAATATA
TG

Human SYNE1 mRNA sequence - var13 (public gi: 16553949) (SEQ ID NO: 195)
ATAGTAGAATTATTCATTATAATATTTGGCTTTGACAAAAATCAGTCTGATCTCGGGAAACCTGGAGAAA
TTTATTTTCTGTACTCTAATGTTCTTTCATTTTGGTGACCATCAAGGTGCTGGGAGAGGAATTAGATGGC
TGTAATTCAAAGTTAATGGAATTAGATGCAGCAGTACAGAAATTCTTGGAACAGAATGGCCAACTGGGTA
AGCCACTGGCCAAGAAGATAGGAAAACTGACTGAACTTCACCAGCAGACCATTAGACAAGCTGAGAATCG
GCTCTCCAAGCTCAATCAGGCAGCATCACATTTAGAAGAATACAATGAAATGCTTGAATTAATTTTGAAG
TGGATTGAAAAAGCTAAAGTCTTGGCTCATGGAACTATTGCATGGAATTCTGCAAGCCAGCTTCGGGAAC
AATATATTTTGCATCAGACCCTGCTAGAAGAATCCAAAGAAATTGACAGTGAGCTGGAAGCAATGACTGA
GAAATTACAGTACCTCACTAGCGTGTACTGTACAGAAAAAATGTCTCAGCAAGTGGCAGAACTGGGACGG
GAGACTGAGGAGTTGCGACAGATGATCAAAATTCGTTTGCAGAACCTCCAAGATGCAGCTAAGGATATGA
AAAAATTTGAAGCAGAGTTGAAAAAGTTACAAGCTGCCTTGGAGCAAGCCCAGGCAACACTGACTTCTCC
AGAAGTTGGACGTCTCAGTCTCAAGGAGCAGCTCTCTCATCGGCAGCATTTGTTGTCTGAGATGGAGTCA
CTGAAGCCGAAGGTGCAAGCAGTGCAGCTCTGCCAGAGTGCCCTCCGGATCCCCGAGGATGTGGTTGCCA
GCTTACCTCTCTGTCATGCTGCTCTGCGGCTGCAGGAAGAGGCCAGCCGGCTGCAGCACACCGCCATCCA
GCAGTGTAACATCATGCAGGAAGCTGTGGTACAATATGAACAATATGAGCAAGAAATGAAACATCTCCAG
CAACTGATAGAAGGAGCTCACAGAGAGATTGAGGATAAACCTGTTGCCACCAGTAACATACAGGAGCTGC
AGGCTCAGATACACGAATGGCCCTCCTCCACTCTGAACTAAGCAGATGCCATCTGCAGAAGTGCTGGTAG

CATAAGGAGGATCGGGTCATAAGCAATCCCAAACTACCAACAAGAGGACCTTGATCTTGGCGAAAGCCCT
CGGTGTGGCAGCTTTAGCCCTCCTCCAGATCACATGTGTGCAAATTATGGCTTCAGAGGTGGAAGATAAA
CAGTGACGGGGGAACAAACAGACAACAAGAAGGTTTGGAAGAAATCTGGTTTGGGACTCTGAACCTTAGC
ACTAAGGAGATTGAGTAAGGACCTCCAAAGTTCCCCGGACTCATGAATTCTGGGCCCTTGGCCCATTCTG
TGCACAGCCAAGGACTTCAGTAGACCATCTGGGCAGCTTTCCCATGGTGCTGCTCCAACCATCAGATAAA
TGACCCTCCCAAGCACCATGTCAGTGTCGTACAATCTACCAACCAACCAGTGCTGAAGAGATTTTAGAAC
CTTGTAACATACAATTTTTAAGAGCTTATATGGCAGCTTCCTTTTTACCTTGTTTTCCTTTGGGGCATGA
TGTTTTAACCTTTGCTTTAGAAGCACAAGCTGTAAATCTAAAAGGCACTTTTTTTTAGAGGTATAAAGAA
AAACTAGATGTAATAAATAAGATCATGGAAGGCTTTATGTGAAAAAAGTTGAATGTTATAGTAAAAAAAA
AAGATATTTATGTATGTACAGTTTGCTAAAGCCAAGTTTTGTTTGTATTGATTTCTTTGCATTTATTATA
GATATTATAAAAT

Human SYNE1 mRNA sequence - var14 (public gi: 12698056) (SEQ ID NO: 196)
ACAACGGAAACTTGAGCAACATAAGGATCTGCTTCAAAACACGGATGCCCACAAAAGAGCATTCCATGAA
ATCTACCGGACCAGGTCTGTTAACGGGATTCCAGTGCCACCTGATCAATTAGAGGACATGGCCGAGAGGT
TTCATTTTGTTTCCTCCACATCAGAGCTACACCTAATGAAAATGGAATTTTTAGAATTAAAGTACCGTCT
GCTCTCACTGCTGGTTCTTGCAGAGTCAAAGCTGAAGTCTTGGATCATTAAGTACGGGAGGAGAGAGTCA
GTGGAGCAGCTTCTACAAAACTACGTGTCTTTTATAGAAAATAGCAAGTTCTTTGAACAATATGAGGTGA
CATACCAGATCTTGAAACAGACAGCTGAGATGTATGTCAAAGCAGATGGTTCAGTGGAAGAAGCTGAGAA
TGTGATGAAATTCATGAATGAAACCACCGCTCAGTGGAGGAATCTCTCAGTAGAAGTGAGGAGTGTGAGG
AGCATGCTGGAAGAAGTGATCTCTAACTGGGATCGCTATGGCAATACAGTGGCTAGTCTGCAAGCCTGGC
TAGAGGATGCTGAAAAAATGCTCAATCAATCAGAAAATGCCAAAAAGGATTTTTTTCGAAATTTACCTCA
TTGGATTCAGCAGCATACTGCCATGAACGATGCTGGCAATTTTCTAATTGAAACCTGTGATGAGATGGTT
TCCCGTGACCTGAAGCAGCAATTACTGTTGCTAAATGGGCGGTGGAGGGAGTTGTTTATGGAAGTCAAGC
AATATGCTCAAGCTGATGAGATGGACAGAATGAAGAAGGAATACACAGACTGTGTTGTTACCCTGTCTGC
TTTTGCAACGGAAGCCCATAAGAAACTTTCTGAACCCTTAGAAGTCTCTTTTATGAATGTCAAGCTATTA
ATTCAAGACTTGGAGGATATTGAGCAGAGGGTGCCTGTGATGGATGCCCAATACAAGATAATTACAAAGA
CAGCACACCTCATTACCAAAGAAAGCCCCCAAGAAGAAGGAAAAGAAATGTTTGCGACCATGTCAAAGCT
CAAAGAGCAGCTAACCAAGGTCAAAGAATGTTACTCCCCACTCCTTTATGAGTCTCAGCAGCTGTTGATT
CCGTTGGAGGAATTAGAAAAGCAGATGACGTCCTTTTATGACTCACTTGGAAAAATCAATGAAATTATCA
CAGTTCTTGAGCGTGAGGCACAATCGAGTGCCCTTTTTAAACAAAAACATCAGGAACTGTTAGCTTGTCA
AGAAAACTGTAAGAAAACCTTGACACTTATTGAGAAAGGCAGTCAAAGTGTTCAAAAGTTTGTGACCTTG
AGCAACGTGTTAAAGCATTTTGATCAGACGAGGCTACAAAGACAGATTGCAGATATTCATGTTGCTTTTC
AGAGTATGGTAAAGAAAACTGGAGATTGGAAGAAGCATGTGGAAACCAACAGTCGCTTGATGAAGAAGTT
TGAGGAGTCTCGAGCAGAGTTGGAGAAGGTACTGCGGATTGCTCAGGAGGGCCTGGAGGAAAAGGGGGAT
CCAGAGGAGCTCCTGCGGAGACACACTGAGTTTTTCAGTCAGCTGGATCAGAGGGTGCTCAATGCTTTCC
TGAAAGCTTGTGATGAACTCACCGACATCCTTCCAGAGCAGGAGCAGCAGGGGCTGCAGGAAGCTGTTCG
AAAGCTCCACAAACAATGGAAGGATCTTCAAGGAGAAGCCCCTTATCATTTGCTTCATCTGAAGATTGAT
GTGGAGAAGAATAGGTTCTTAGCCTCTGTAGAAGAATGCAGAACTGAGCTGGATCGAGAGACCAAGCTGA
TGCCCCAGGAAGGCAGTGAAAAGATAATTAAAGAGCACAGGGTTTTCTTCAGTGACAAAGGTCCTCATCA
TCTCTGTGAGAAAAGGTTACAGCTCATCGAGGAACTCTGTGTGAAACTCCCAGTGCGGGACCCAGTAAGG
GACACACCTGGAACCTGTCACGTGACTCTCAAAGAGCTCAGAGCTGCCATTGACAGCACCTACAGGAAGC
TCATGGAAGACCCAGACAAGTGGAAGGACTACACTAGCAGATTCTCTGAGTTCTCATCTTGGATATCTAC
AAATGAGACACAATTAAAGGGGATCAAGGGTGAGGCCATCGATACTGCCAACCACGGAGAGGTTAAACGT
GCCGTTGAAGAGATCAGAAATGGTGTTACCAAAAGGGGTGAGACCCTCAGCTGGCTGAAATCCAGGCTGA
AAGTTTTGACAGAAGTTTCTTCTGAGAATGAAGCCCAAAAGCAGGGAGATGAGCTGGCAAAATTATCCAG
CTCTTTCAAGGCTCTTGTGACGCTGCTGTCAGAGGTTGAAAAGATGCTAAGCAATTTTGGGGACTGTGTC
CAGTACAAAGAAATAGTCAAAAATTCTCTCGAAGAATTAATTTCTGGCTCTAAAGAAGTCCAGGAACAAG
CTGAGAAGATCTTGGATACTGAAAATCTGTTTGAAGCACAGCAGTTACTTCTTCATCACCAGCAAAAGAC
AAAGCGGATCTCAGCAAAGAAGAGAGATGTGCAGCAGCAGATCGCGCAGGCGCAGCAGGGAGAAGGGGGG
CTGCCTGACCGAGGCCACGAGGAGCTGCGGAAGCTGGAGAGCACACTGGATGGCCTGGAGCGCAGCCGGG
AGAGGCAGGAACGCCGCATCCAGGTCACATTAAGAAAATGGGAGCGATTTGAAACAAACAAAGAAACAGT
AGTAAGATACCTTTTTCAAACAGGTTCCAGTCATGAACGCTTCTTGAGTTTTAGCAGTTTGGAAAGTTTA
TCTTCAGAACTGGAACAAACAAAGGAGTTTTCTAAACGGACAGAAAGTATTGCAGTCCAGGCTGAGAACC
TTGTAAAGGAAGCTTCAGAGATACCGCTTGGGCCCCAAAATAAGCAGCTGCTTCAACAGCAGGCCAAGTC
AATCAAAGAACAAGTCAAAAAATTAGAAGACACGCTTGAAGAAGAGTATGTGATTGACAAGTCCTAAACT
TTCTTCTCTGAGATAAAGTTTCATACAATCTTTCCTGTACCTTGTATTCAAAACACTCTTTTAAAATCTC
AAAGTGTCTGTGTATTTCAGCATGTTTTGAGGAAACAACTCACAGTTCAAAAGAAAGTATCGCTAATACA
GAAACCAATATCTATAACAGAGCCCAAAAAAATATAAAGGATGTGGGTTTTGCATCTTAAACTGATCATGT
TCATGAGAAAGCCATATCTATTCTATTCTGTGGCCTTTGTACATTGTAGAGGGAATCTTGAAAAAGAACT
AATATTTAAAATAATTTTTTTACTATATTATTCTGCTGTCACCATTTAGAGCGAAAAGGAGATATTTTGT
TAGTGTAGATTCCAGGCCTAAATACACATCACATAGACCATATATCTCCAACCTGAAGAAGCTCCTGGAG
CTTGTTTACAGTGCCTCGGTATTCAAGTTATCCTGACTAATATGCTCTTTCCAGAAATTAACTTTAAAAT

ATTTTATTTTTAAACTTTTAATGTTTGTTTATCTG

Human SYNE1 mRNA sequence - var15 (public gi: 2895592) (SEQ ID NO: 197)
CAACCTGCATAGTAACGAAACCAAACGGCTGGTGTGATTGACCGATGGGAGCTTCTCCAGGCCCAGGCAT
TGAGCAAGGAGTTGAGGATGAAGCAGAACCTCCAGAAGTGGCAGCAGTTTAACTCAGACTTGAACAGCAT
CTGGGCCTGGCTGGGGGACACGGAGGAGGAGTTGGAACAGCTCCAGCGTCTGGAACTCAGCACTGACATC
CAGACCATCGAGCTCCAGATCAAAAAGCTCAAGGAGCTCCAGAAAGCTGTGGACCACCGCAAAGCCATCA
TCCTCTCCATCAATCTCTGCAGCCCTGAGTTCACCCAGGCTGACAGCAAGGAGAGCCGGGACCTGCAGGA
TCGCTTGTCGCAGATGAATGGGCGCTGGGACCGAGTGTGCTCTCTGCTGGAGGAGTGGCGGGGCCTGCTG
CAGGATGCCCTGATGCAGTGCCAGGGTTTCCATGAAATGAGCCATGGTTTGCTTCTTATGCTGGAGAACA
TTGACAGAAGGAAAAATGAAATTGTCCCTATTGATTCTAACCTTGATGCAGAGATACTTCAGGACCATCA
CAAACAGCTTATGCAAATAAAGCATGAGCTGTTGGAATCCCAACTCAGAGTAGCCTCTTTGCAAGACATG
TCTTGCCAACTACTGGTGAATGCTGAAGGAACAGACTGTTTAGAAGCCAAAGAAAAAGTCCATGTTATTG
GAAATCGGCTCAAACTTCTCTTGAAGGAGGTCAGTCGTCATATCAAGGAACTGGAGAAGTTATTAGACGT
GTCAAGTAGTCAGCAGGATTTGTCTTCCTGGTCTTCTGCTGATGAACTGGACACCTCAGGGTCTGTGAGT
CCCACATCAGGAAGGAGCACCCCAAACAGACAGAAAACGCCACGAGGCAAGTGTAGTCTCTCACAGCCTG
GACCCTCTGTCAGCAGTCCACATAGCAGGTCCACAAAAGGTGGCTCCGATTCCTCCCTTTCTGAGCCAGG
GCCAGGTCGGTCCGGCCGCGGCTTCCTGTTCAGAGTCCTCCGAGCAGCTCTTCCCCTTCAGCTTCTCCTG
CTCCTCCTCATCGGGCTTGCCTGCCTTGTACCAATGTCAGAGGAAGACTACAGCTGTGCCCTCTCCAACA
ACTTTGCCCGGTCATCCACCCCATGCTCAGATACACGAATGGCCCTCCTCCACTCTGAACTAAGCAGATG
CCATCTGCAGAAGTGCTGGTAGCATAAGGAGGATCGGGTCATAAGCAATCCCAAACTACCAACAAGAGGA
CCTTGATCTTGGCGAAAGCCATCGGTGTGGCAGCTTTAGCCCTCCTCCAGATCACATGTGTGCAAATTAT
GGCTTCAGAGGGTGGAAGATAAACAGTGACGGGGGAACAAACAGACAACAAGAAGGTTTGGAAGAAATCT
GGTTTGAGACTCTGAACCTTAGCACTAAGGAGATTGAGTAAGGACCTCCAAAGTTCCCCGGACTCATGAA
TTCTGGGCCCTTGGCATTCGTGTGCACAGCCAAGGACTTCAGTAGACCATCTGGGCAGCTTTCCCATGGT
GCTGCTCCAACCATCAGATAAATGACCCTCCCCAAGCACCATGTCAGTGTCGTACAATCTACCAACCAAC
CAGTGCTGAAGAGATTTTAGAACCTTGTAACATACAATTTTTAAGAGCTTATATGGCAGCTTCCTTTTTA
CCTTGTTTTCCTTTGGGGCATGATGTTTTAACCTTTTGCTTTAGAAGCACAAGCTGTAAATCTAAAAGGC
ACTTTTTTTTAGAGGTATAAGAAAAACTAGATGTAATAAATAAGATCATGGAAGGCTTTATGTGAAAAAA
GTTGAATGTTATAGTAAAAAAAAAAA

Human SYNE1 mRNA sequence - var16 (public gi: 6330956) (SEQ ID NO: 198)
CTCGATTTGTGCCGTCAGTCTAACAACCTGTGCTTGCAAAGGGAAGAGGATCTTCAGAGAACAAGAGATT
ACCATGACTGTATGAATGTTGTTGAAGTGTTCCTAGAAAAATTTACTACAGAATGGGATAACTTGGCCAG
ATCTGATGCAGAGAGTACAGCTGTCCACCTGGAAGCTTTGAAAAAGTTAGCATTGGCATTGCAGGAGAGA
AAGTATGCTATTGAAGATCTGAAAGATCAAAAGCAGAAAATGATAGAGCATCTGAATTTAGATGACAAGG
AGTTAGTCAAAGAACAGACGAGTCATTTAGAGCAACGTTGGTTTCAGCTTGAGGACCTCATTAAAAGGAA
AATCCAAGTGTCAGTCACCAACTTGGAGGAGTTAAATGTGGTGCAGTCCAGATTTCAGGAGCTAATGGAG
TGGGCAGAAGAGCAACAACCCAACATCGCCGAGGCCCTTAAGCAGAGCCCTCCTCCAGATATGGCTCAGA
ACCTTCTCATGGATCACCTGGCCATCTGCAGTGAACTGGAGGCCAAGCAGATGCTCCTGAAATCGCTTAT
AAAGGACGCAGACAGGGTCATGGCAGATCTTGGTCTCAATGAGCGACAGGTCATCCAGAAGGCTCTCTCT
GATGCACAAAGCCACGTGAATTGTCTCAGTGACTTAGTGGGCCAGCGAAGAAAGTACTTAAACAAAGCCT
TGTCCGAGAAAACCCAGTTTCTCATGGCAGTGTTCCAGGCCACCAGCCAAATTCAGCAACATGAGCGAAA
GATAATGTTCCGTGAACACATCTGTCTGTTACCAGATGATGTGAGCAAACAAGTCAAAACATGTAAGAGT
GCACAAGCCAGCCTCAAGACTTACCAAAATGAAGTCACTGGACTTTGGGCCCAGGGTCGCGAACTAATGA
AGGAAGTCACAGAGCAGGAAAAGAGTGAAGTGCTGGGGAAGCTTCAGGAATTGCAGAGTGTCTATGACAG
TGTTTTACAAAAGTGCAGTCACCGGTTACAAGAACTAGAGAAGAATTTGGTTTCTAGGAAGCATTTTAAG
GAAGATTTTGATAAAGCTTGCCACTGGCTAAAACAAGCAGATATTGTTACATTTCCTGAAATCAACCTAA
TGAATGAGAGTACTGAGCTTCATACACAACTGGCTAAATACCAAAACATTCTTGAACAATCTCCAGAATA
TGAAAATCTTCTACTTACGCTGCAGAGAACTGGGCAGACCATATTACCATCGCTGAATGAAGTCGATCAT
TCCTACCTCAGTGAAAAGCTAAATGCTTTGCCTCGACAATTTAATGTAATTGTTGCCTTGGCTAAAGACA
AGTTCTATAAAGTCCAGGAAGCAATTCTTGCTCGGAAGGAATATGCTTCCTTGATTGAGTTGACAACCCA
GTCTCTCAGTGAACTTGAAGCCCAATTCTTGAGGATGAGCAAAGTTCCCACCGACCTGGCCGTTGAGGAG
GCTCTTTCTCTGCAAGATGGTTGCAGAGCCATTCTGGACGAGGTGGCGGGCCTTGGGGAGGCGGTGGATG
AACTGAACCAGAAAAAAGAAGGTTTTCGCAGCACAGGTCAGCCTTGGCAGCCAGACAAGATGCTGCACCT
TGTCACCTTATATCACAGGCTGAAGCGACAAACAGAACAGAGGGTTAGCTTATTAGAAGACACCACCAGT
GCTTACCAAGAACACGAGAAGATGTGCCAACAGCTGGAGAGACAACTGAAGTCTGTAAAAGAGGAGCAGT
CCAAAGTGAATGAGGAAACGCTGCCTGCAGAGGAGAAGCTCAAAATGTATCACTCCCTGGCAGGAAGTCT
CCAGGACTCAGGGATTGTACTGAAACGAGTAACCATACATCTTGAAGATCTTGCCCCACACCTTGACCCC
TTGGCTTATGAGAAAGCCAGGCATCAGATCCAGTCCTGGCAAGGGGAGTTAAAACTGTTGACTTCTGCCA
TTGGTGAGACGGTGACAGAATGTGAGAGCCGAATGGTGCAGAGTATAGACTTCCAGACTGAGATGAGTCG
CTCCCTGGACTGGCTGAGGAGAGTGAAGGCAGAGCTCAGTGGGCCGGTGTACCTAGACCTCAACCTGCAG

GACATCCAAGAGGAAATCAGAAAAATCCAAATTCATCAGGAAGAGGTCCAGTCCAGCTTGAGAATCATGA
ATGCGCTGAGTCACAAGGAAAAGGAGAAGTTCACAAAGGCCAAGGAGCTGATTTCTGCGGATTTAGAACA
CAGCCTCGCTGAGCTCTCAGAGCTGGATGGAGACATCCAGGAAGCCTTACGCACCAGACAGGCTACCTTG
ACTGAAATATATAGCCAGTGTCAAAGGTATTATCAGGTATTTCAAGCAGCCAATGACTGGCTTGAGGATG
CCCAAGAAATGTTACAGCTGGCAGGCAATGGCCTAGACGTGGAGAGCGCAGAGGAAAATCTCAAAAGCCA
CATGGAATTTTTCAGTACAGAGGATCAGTTCCATAGTAACCTGGAGGAGCTCCACAGCCTGGTAGCCACC
CTGGACCCACTCATCAAGCCAACCGGCAAAGAAGACCTAGAACAGAAAGTGGCTTCTCTGGAACTCAGGA
GCCAGAGGATGAGCCGGGACTCTGGTGCCCAAGTGGATCTCTTGCAGAGATGCACAGCTCAATGGCACGA
TTACCAGAAAGCAAGGGAAGAGGTTATTGAATTGATGAATGATACAGAAAAGAAATTGTCTGAGTTTTCT
TTGTTGAAGACTTCGTCTAGTCATGAAGCGGAAGAAAAATTGTCAGAACACAAGGCTTTAGTGTCAGTGG
TTAACTCTTTCCATGAGAAAATTGTGGCCCTTGAGGAAAAAGCTTCACAACTGGAGAAAACCGGAAATGA
TGCCAGCAAAGCCACCCTGAGCAGGTCAATGACCACCGTCTGGCAGCGCTGGACACGCCTTCGAGCTGTG
GCCCAGGACCAGGAGAAGATCCTGGAAGATGCAGTGGATGAGTGGACGGGCTTTAACAACAAGGTTAAAA
AGGCCACTGAAATGATTGATCAGCTGCAAGATAAGTTACCTGGAAGTTCAGCAGAGAAAGCATCGAAAGC
AGAGCTCTTAACTCTTCTTGAATACCACGACACGTTCGTTCTGGAGCTGGAGCAGCAGCAGTCGGCCTTG
GGCATGCTGCGGCAGCAAACCCTGAGCATGCTCCAGGATGGAGCCGCCCCAACCCCTGGGGAAGAGCCTC
CGCTCATGCAGGAAATCACCGCCATGCAAGATCGGTGCCTGAACATGCAGGAGAAAGTGAAGACTAATGG
AAAGTTGGTGAAGCAAGAGCTGAAGGACCGAGAAATGGTGGAGACTCAGATCAATTCTGTGAAATGTTGG
GTTCAGGAAACGAAAGAATATTTAGGGAATCCAACAATAGAAATAGATGCTCAACTTGAAGAACTTCAGA
TTCTCCTAACAGAAGCCACAAATCACCGACAGAACATTGAAAAAATGGCAGAAGAACAGAAGGAGAAGTA
CTTAGGTCTTTATACCATATTACCTTCTGAACTCTCCCTTCAGTTGGCTGAAGTGGCGTTAGATCTAAAG
ATCCGAGATCAGATCCAAGACAAAATAAAAGAAGTTGAGCAGAGCAAGGCCACGAGCCAGGAACTCAGCC
GGCAAATTCAGAAGTTAGCTAAAGACCTCACAACTATTCTAACTAAGCTGAAAGCGAAGACAGATAATGT
AGTTCAAGCTAAAACTGACCAAAAGGTGCTGGGAGAGGAATTAGATGGCTGTAATTCAAAGTTAATGGAA
TTAGATGCAGCAGTACAGAAATTCTTGGAACAGAATGGCCAACTGGGTAAGCCACTGGCCAAGAAGATAG
GAAAACTGACTGAACTTCACCAGCAGACCATTAGACAAGCTGAGAATCGGCTCTCCAAGCTCAATCAGGC
AGCATCACATTTAGAAGAATACAATGAAATGCTTGAATTAATTTTGAAGTGGATTGAAAAAGCTAAAGTC
TTGGCTCATGGAACTATTGCATGGAATTCTGCAAGCCAGCTTCGGGAACAATATATTTTGCATCAGGTAA
CCTTAGGAAAAATAATCTTTAAAAAGTAACCAAGGGCAATTTGATTTAACTGGGTAGACTGACACAACAC
TTAGAGGGCTGTGATGTAAAATTTTTGGAGCTACCAGATAAAAAGAATGCTAAGGTACCCCTAAGTTGTT
CAGTAGTTGGACAGAAAGGAGCTTCTCATGAAATTTCATGAAATTGAATAAATAAATATCCTTGATCTTC
CCTAAACCTACCTTACACCAAGACCCAAACCAATCAGCCTTGTAGAACTCATTTTCTGTAGCTTCTTTGA
AATAATTATCTGCAGGGATCTGGTGGGAAATTCTTTCCTGTGAAGAGATGCAATGAAGTGTGGAAAGATT
CTAGACTCCACACTCAGACTGGTGGGAAAACCAACCTCCGCCATGCAGGGCTGTGTGATTTGGAGCAGAA
TGCTTTGCCTCTCTGAATTCTGTCTTCTCATCATTTGTATGAAGACGTAAATAATATTCGTATTTCAGAC
TTATGAGATCAAGTGGTTTAAGGTACACACGTGCAAACGTCTGCCTGGCACATGCAGGTGCTCAGTGGGA
GATCTCCCGCCTCCTCCCTCAGCCCTCACCCAGGCCTGTCATCTGGCCTTCCACAGGAGGTCGGGCAGCC
CAGAGCAAGCCATGAGTCCACATCACATGCTGGCTATGTTAGTTCATTTCCTCTGAAGTTACATGAGAAA
AATGTTCCTTTTCTGTCAGTCACGTCATCCAGGAAATTATTTCATCCTTTTGTAACTTAAGCTTAAATTA
GACACAGATAGTTAATAGGCTAGTTATCATATAATAAAATATAGGGTGACTTTTATAGGAGTTACATGGG
TATCGAGTATTCTAGATTTTTGTCTCTTATATTATTTATGTATCCTTGTGGCCTTTAAATGAATCCCTGT
TTCCATTCTTGTTTACAGGGTTCTAGATCAAAGCCTCATTTTTCCATTTTTGGAATGCTTTAACAGCTTC
TAATTTTTCCCTATATTCACAGTCTTTCTTCTCTGATCAATCTTGCGTATTCTTCCCACAATGTCTTTCT
TAAGCAACTCCAATCTTTTGCTTTAAGATATGCTTAGATATGAACAGACAGGACTTAAGTTACCACTGAT
TTGAAAACAATGAAAAAAAGCCAACATCCTTAGAAGTCTAGAAATGCAAATTTCAGCAAAAAAAGAGAGG
AAGAAAGACAAACTTAACTGTCACATTCATACTGTTTCTTTCAAGTTCATATTTAAGGAAGTGAGAGCTC
TCAAACATTGCTGGTATCCTGGTAAAATCTCTTTGAAAAATAATTGGCAAAATGTATGGTGATTGTCAAA
AATGTTGCTACTCTGGGCCACGTGCGGTGGCTCACACCTGTAGTCCCAGCACTTTGGGAGGCCGAGGTGG
GTGGATCACAGGGTCAGGAGTTCGGGACCGGCTTGGCCGGTATGGTGAAGCCCCATCTCTACTAGGAGTG
CAAAAGTCAGCTGGGCGTGGTGGTGGGCGCCTGTAGTCCCAGCTACTCGGGAATCTGAGGCGGGAGAATC
GCTTGAACTCGGGAGGTGGAGGTTGCAGTGAGCCAAGATCATGCCACTGCACTCCAGCCTGGGTGACAGT
GAGACTCCATCTC

Human SYNE1 mRNA sequence - var17 (public gi: 20521661) (SEQ ID NO: 199)
GTTGGATTTCTCTAATGGAAAATGTTATTCAGAAGGATGAAGATAATATTAAAAATTCCATAGGTTACAA
GGCAATTCATGAATACCTTCAGAAATATAAGGGTTTTAAGATAGACATTAACTGTAAACAGCTGACAGTG
GATTTTGTGAACCAGTCCGTGCTACAAATCAGCAGTCAGGATGTGGAAAGTAAGCGTAGTGATAAGACTG
ATTTTGCTGAGCAACTTGGAGCAATGAATAAAAGTTGGCAAATTCTGCAAGGTCTAGTAACTGAGAAGAT
CCAGCTGTTGGAAGGCTTATTGGAATCTTGGTCAGAATATGAAAATAATGTACAATGTCTGAAAACATGG
TTTGAAACCCAGGAAAAGAGACTAAAACAACAGCATCGAATTGGAGATCAGGCTTCTGTTCAAAATGCAC
TGAAAGACTGTCAGGATCTGGAAGATTTGATTAAAGCAAAAGAAAAAGAAGTAGAGAAAATTGAGCAGAA
TGGACTTGCTTTGATTCAGAACAAGAAAGAAGACGTCTCTAGCATTGTCATGAGCACACTGCGAGAGCTC
GGCCAAACCTGGGCAAATTTAGATCACATGGTTGGACAATTAAAGATACTGCTGAAATCAGTGCTTGACC

AATGGAGTAGTCACAAAGTGGCCTTTGACAAGATAAACAGTTACCTCATGGAGGCCAGATACTCTCTTTC
CCGATTCCGTCTGCTGACTGGCTCCTTAGAAGCTGTGCAAGTTCAGGTGGACAATCTTCAGAATCTCCAA
GATGATCTGGAAAAACAGGAAAGGAGCTTACAGAAATTTGGCTCTATCACCAACCAATTATTAAAAGAGT
GTCACCCACCCGTGACAGAAACTCTTACCAATACACTGAAAGAAGTCAACATGAGATGGAATAACTTGCT
GGAAGAGATTGCTGAGCAGCTACAGTCCAGCAAGGCCCTACTTCAGCTTTGGCAAAGATACAAGGACTAC
TCCAAACAGTGTGCTTCGACAGTTCAGCAGCAGGAGGATCGAACCAATGAGCTGTTGAAGGCAGCCACAA
ACAAGGACATTGCCGATGATGAGGTTGCCACATGGATTCAAGATTGCAACGACCTCCTCAAAGGACTGGG
CACAGTTAAAGATTCCCTCTTTGTTCTCCATGAGCTGGGAGAGCAACTGAAGCAACAAGTGGATGCTTCC
GCAGCATCAGCTATTCAATCGGATCAACTCTCTTTGAGTCAACACTTGTGTGCCCTGGAGCAAGCTCTCT
GCAAACAGCAGACTTCATTACAGGCTGGAGTTCTTGATTATGAAACCTTTGCCAAGAGTTTAGAAGCTTT
GGAGGCCTGGATAGTGGAAGCTGAAGAAATACTACAAGGGCAGGACCCTAGCCACTCATCTGACCTCTCC
ACAATCCAGGAAAGGATGGAAGAACTTAAGGGACAGATGTTAAAATTCAGCAGCATGGCTCCAGATTTAG
ACCGTCTAAATGAGCTTGGATATAGGTTACCCTTGAATGATAAGGAAATCAAAAGAATGCAGAATCTGAA
CCGCCATTGGTCTCTGATCTCCTCTCAGACTACAGAAAGATTCAGCAAGTTGCAGTCATTTTTGCTACAA
CATCAGACTTTCTTGGAAAAAATGTGAAACATGGATGGAATTCCTAGTTCAGACAGAACAAAAGTTAGCAG
TAGAGATTTCAGGAAATTATCAGCACCTTTTGGAACAGCAGAGAGCACACGAGTTGTTTCAAGCCGAGAT
GTTCAGTCGTCAGCAGATTTTGCACTCAATCATTATTGATGGGCAACGTCTTCTAGAACAAGGTCAAGTT
GATGACAGGGATGAATTCAACCTGAAATTGACACTCCTCAGTAATCAATGGCAGGGAGTGATTCGCAGGG
CCCAGCAGAGGCGGGGGATCATTGACAGCCAGATTCGCCAGTGGCAGCGCTATAGGGAGATGGCAGAAAA
GCTTCGTAAATGGTTGGTTGAAGTGTCCTACCTCCCCATGAGTGGTCTCGGAAGTGTTCCTATACCACTG
CAACAAGCAAGGACCCTCTTTGATGAAGTGCAGTTCAAAGAAAAAGTGTTTCTGCGGCAACAAGGCAGCT
ACATCCTGACTGTGGAGGCTGGCAAGCAACTCCTTCTCTCGGCGGACAGTGGCGCTGAGGCCGCCTTGCA
GGCCGAACTCGCTGAAATCCAAGAGAAATGGAAATCAGCCAGCATGCGGCTGGAAGAACAGAAGAAAAAA
CTAGCCTTCTTGTTGAAAGACTGGGAAAAATGTGAGAAAGGAATAGCAGATTCCCTGGAGAAACTACGAA
CTTTCAAAAAGAAGCTTTCGCAGTCTCTCCCGGATCACCATGAAGAGCTCCATGCAGAACAAATGCGTTG
CAAGGAATTAGAAAATGCAGTTGGGAGCTGGACAGATGACTTGACCCAGTTGAGCCTGCTGAAGGACACC
CTCTCTGCCTATATCAGTGCTGATGATATCTCCATTCTTAATGAACGCGTAGAGCTTCTGCAAAGGCAGT
GGGAAGAACTATGCCACCAGCTCTCCTTAAGGCGGCAGCAAATAGGTGAAAGATTGAATGAATGGGCAGT
CTTCAGTGAAAAGAACAAGGAACTCTGTGAGTGGTTGACTCAAATGGAAAGCAAAGTTTCTCAGAATGGA
GACATTCTCATTGAAGAAATGATAGAGAAGCTCAAGAAGGATTATCAAGAGGAAATTGCTATTGCTCAAG
AGAACAAAATACAGCTCCAACAAATGGGAGAACGACTTGCTAAAGCCAGCCATGAAAGCAAAGCATCTGA
GATTGAATACAAGCTGGGAAAGGTCAACGACCGGTGGCAGCATCTCCTGGACCTCATTGCAGCCAGGGTG
AAGAAGCTGAAGGAGACCCTGGTAGCCGTGCAGCAGCTTGATAAGAACATGAGCAGCCTGAGGACCTGGC
TCGCTCACATCGAGTCAGAGCTGGCCAAGCCAATAGTCTACGATTCCTGTAACTCGGAAGAAATACAGAG
AAAGCTTAATGAGCAGCAGGAGCTTCAGAGAGACATAGAGAAGCACAGTACAGGTGTTGCATCTGTCCTC
AACCTGTGTGAAGTCCTGCTGCACGACTGTGACGCCTGTGCCACTGATGCCGAGTGTGACTCTATACAGC
AGGCTACGAGAAACCTGGACCGGCGGTGGAGAAACATTTGTGCTATGTCCATGGAAAGGAGGCTGAAAAT
CGAAGAGACGTGGCGATTGTGGCAGAAATTTCTGGATGACTATTCACGTTTTGAAGATTGGCTGAAGTCT
TCAGAAAGGACAGCTGCTTTTCCCAGCTCTTCTGGGGTGATCTATACAGTTGCCAAGGAAGAACTAAAGA
AATTTGAGGCTTTCCAGCGACAGGTCCACGAGTGCCTGACGCAGCTGGAACTGATCAACAAGCAGTACCG
CCGCCTGGCCAGGGAGAACCGCACTGATTCAGCATGTAGCCTCAAACAGATGGTTCACGAAGGCAACCAG
AGATGGGACAACCTGCAAAAGCGTGTCACCTCCATCTTGCGCAGACTCAAGCATTTTATTGGCCAGCGTG
AGGAGTTTGAGACTGCGCGGGACAGCATTCTGGTCTGGCTCACAGAGATGGATCTGCAGCTCACTAATAT
TGAACATTTTTCTGAGTGTGATGTTCAAGCTAAAATAAAGCAACTCAAGGCCTTCCAGCAGGAAATTTCA
CTGAACCACAATAAGATTGAGCAGATAATTGCCCAAGGAGAACAGCTGATAGAAAAGAGTGAGCCCTTGG
ATGCAGCGATCATCGAGGAGGAACTAGATGAGCTCCGACGGTACTGCCAGGAGGTCTTCGGGCGTGTGGA
AAGATACCATAAGAAACTGATCCGCCTGCCTCTCCCAGACGATGAGCACGACCTCTCAGACAGGGAGCTG
GAGCTGGAAGACTCTGCAGCTCTGTCGGACCTGCACTGGCACGACCGCTCTGCAGACAGCCTGCTTTCTC
CACAGCCTTCCTCCAATCTCTCCCTCTCGCTCGCTCAGCCCCTCCGGAGCGAGCGGTCAGGACGAGACAC
CCCAGCTAGTGTGGACTCCATCCCCCTGGAGTGGGATCACGACTATGACCTCAGTCGGGACCTGGAGTCT
GCAATGTCCAGAGCTCTGCCCTCTGAGGATGAAGAAGGTCAGGATGACAAAGATTTCTACCTCCGGGGAG
CTGTTGCCTTATCAGGGGACCACAGTGCCCTAGAGTCACAGATCCGACAACTGGGCAAAGCCCTGGATGA
TAGCCGCTTTCAGATACAGCAAACCGAAAATATCATTCGCAGCAAAACTCCCACGGGGCCGGAGCTAGAC
ACCAGCTACAAAGGCTACATGAAACTGCTGGGCGAATGCAGTAGCAGTATAGACTCCGTGAAGAGACTGG
AGCACAAACTGAAGGAGGAAGAGGAGAGCCTTCCTGGCTTTGTTAACCTGCATAGTACCGAAACCCAAAC
GGCTGGTGTGATTGACCGATGGGAGCTTCTCCAGGCCCAGGCATTGAGCAAGGAGTTGAGGATGAAGCAG
AACCTCCAGAAGTGGCAGCAGTTTAACTCAGACTTGAACAGCATCTGGGCCTGGCTGGGGGACACGGAGG
AGGAGTTGGAACAGCTCCAGCGTCTGGAACTCAGCACTGACATCCAGACCATCGAGCTCCAGATCAAAAA
GCTCAAGGAGCTCCAGAAAGCTGTGGACCACCGCAAAGCCATCATCCTCTCCATCAATCTCTGCAGCCCT
GAGTTCACCCAGGCTGACAGCAAGGAGAGCCGGGACCTGCAGGATCGCTTGTCGCAGATGAATGGGCGCT
GGGACCGAGTGTGCTCTCTGCTGGAGGAGTGGCGGGGCCTGCTGCAGGATGCCCTGATGCAGTGCCAGGG
TTTCCATGAAATGAGCCATGGTTTGCTTCTTATGCTGGAGAACATTGACAGAAGGAAAAATGAAATTGTC
CCTATTGATTCTAACCTTGATGCAGAGATACTTCAGGACCATCACAAACAGCTTATGCAAATAAAGCATG

AGCTGTTGGAATCCCAACTCAGAGTAGCCTCTTTGCAAGACATGTCTTGCCAACTACTGGTGAATGCTGA
AGGAACAGACTGTTTAGAAGCCAAAGAAAAAGTCCATGTTATTGGAAATCGGCTCAAACTTCTCTTGAAG
GAGGTCAGTCGTCATATCAAGGAACTGGAGAAGTTATTAGACGTGTCAAGTAGTCAGCAGGATTTGTCTT
CCTGGTCTTCTGCTGATGAACTGGACACCTCAGGGTCTGTGAGTCCCACATCAGGAAGGAGCACCCCAAA
CAGACAGAAAACGCCACGAGGCAAGTGTAGTCTCTCACAGCCTGGACCCTCTGTCAGCAGTCCACATAGC
AGGTCCACAAAAGGTGGCTCCGATTCCTCCCTTTCTGAGCCAGGGCCAGGTCGGTCCGGCCGCGGCTTCC
TGTTCAGAGTCCTCCGAGCAGCTCTTCCCCTTCAGCTTCTCCTGCTCCTCCTCATCGGGCTTGCCTGCCT
TGTACCAATGTCAGAGGAAGACTACAGCTGTGCCCTCTCCAACAACTTTGCCCGGTCATTCCACCCCATG
CTCAGATACACGAATGGCCCTCCTCCACTCTGAACTAAGCAGATGCCATCTGCAGAAGTGCTGGTAGCAT
AAGGAGGATCGGGTCATAAGCAATCCCAAACTACCAACAAGAGGACCTTGATCTTGGCGAAAGCCCTCGG
TGTGGCAGCTTTAGCCCTCCTCCAGATCACATGTGTGCAAATTATGGCTTCAGAGGTGGAAGATAAACAG
TGACGGGGGAACAAACAGACAACAAGAGGTTTGGAAGAAATCTGGTTTGAGACTCTGAACCTTAGCACT
AAGGAGATTGAGTAAGGACCTCCAAAGTTCCCCGGACTCATGAATTCTGGGCCCTTGGCCCATTCTGTGC
ACAGCCAAGGACTTCAGTAGACCATCTGGGCAGCTTTCCCATGGTGCTGCTCCAACCATCAGATAAATGA
CCCTCCCAAGCACCATGTCAGTGTCGTACAATCTACCAACCAACCAGTGCTGAAGAGATTTTAGAACCTT
GTAACATACAATTTTTAAGAGCTTATATGGCAGCTTCCTTTTTACCTTGTTTTCCTTTGGGGCATGATGT
TTTAACCTTTGCTTTAGAAGCACAAGCTGTAAATCTAAAAGGCACTTTTTTTTAGAGGTATAAAGAAAAA
CTAGATGTAATAAATAAGATCATGGAAGGCTTTATGTGAAAAAAGTTGAATGTTATAGT

Human SYNE1 mRNA sequence - var18 (public gi: 28195688) (SEQ ID NO: 200)
TGTTCTCACGAGGGGGCCAGCTTGGGGCTTGACTGAGCAGGAGCTTCCATGGTCCCACACGTAGTATGAC
ATGTGACCTCTGCACATTGTTTACAGTTCCTAAACTGTGATTTCTTTTTCTGTGAAATAGTTATAATAGT
AAGTGGCTACCAAGTAGAAAGTGGTCATGGGGGGTGAAGGTTAAACACAATAACGGACACACAGAACTTA
CACAGGGCATTTTATGCCAAGCTATATTGAATATCTATATCCCTCTACCTGCCCGTCAATGTCATGAATA
TTGACAATTCACTCTAGACCCTGCTAGAAGAATCCAAAGAAATTGACAGTGAGCTGGAAGCAATGACTGA
GAAATTACAGTACCTCACTAGCGTGTACTGTACAGAAAAAATGTCTCAGCAAGTGGCAGAACTGGGACGG
GAGACTGAGGAGTTGCGACAGATGATCAAAATTCGTTTGCAGAACCTCCAAGATGCAGCTAAGGATATGA
AAAAATTTGAAGCAGAGTTGAAAAAGTTACAAGCTGCCTTGGAGCAAGCCCAGGCAACACTGACTTCTCC
AGAAGTTGGACGTCTCAGTCTCAAGGAGCAGCTCTCTCATCGGCAGCATTTGTTGTCTGAGATGGAGTCA
CTGAAGCCGAAGGTGCAAGCAGTGCAGCTCTGCCAGAGTGCCCTCCGGATCCCCGAGGATGTGGTTGCCA
GCTTACCTCTCTGTCATGCTGCTCTGCGGCTGCAGGAAGAGGCCAGCCGGCTGCAGCACACCGCCATCCA
GCAGTGTAACATCATGCAGGAAGCTGTGGTACAATATGAACAATATGAGCAAGAAATGAAACATCTCCAG
CAACTGATAGAAGGAGCTCACAGAGAGATTGAGGATAAACCTGTTGCCACCAGTAACATACAGGAGCTGC
AGGCTCAGATTTCTCGGCATGAGGAGCTGGCGCAGAAAATTAAGGGCTACCAGGAGCAGATCGCTTCTTT
GAATTCCAAGTGCAAGATGCTGACGATGAAAGCCAAGCACGCCACCATGCTGCTGACGGTGACCGAGGTC
GAGGGGCTGGCGGAAGGGACAGAGGACCTGGATGGGGAGCTCCTCCCCACGCCTTCGGCCCACCCCTCTG
TGGTCATGATGACTGCAGGTCGCTGTCACACTTTGCTGTCACCGGTCACTGAGGAGTCTGGGGAGGAGGG
AACCAACAGTGAGATTTCCTCTCCACCTGCCTGTCGCTCCCCTTCACCTGTGGCTAATACAGATGCTTCT
GTTAACCAGGACATTGCATATTACCAAGCCTTGTCTGCTGAGAGGTTGCAGACAGATGCTGCAAAAATTC
ACCCCAGCACATCCGCATCCCAGGAGTTCTATGAACCGGGATTGGAGCCATCCGCTACTGCCAAACTGGG
TGATTTGCAGCGTTCTTGGGAAACCTTAAAGAATGTGATCAGTGAGAAGCAGCGCACACTCTATGAAGCT
TTGGAGCGCCAGCAGAAGTACCAGGACTCCCTCCAGTCCATCTCTACGAAGATGGAGGCCATTGAGCTGA
AACTCAGTGAGAGCCCAGAGCCTGGCAGGAGTCCAGAAAGCCAGATGGCTGAACATCAGGCATTGATGGA
TGAGATTCTCATGCTCCAGGATGAAATCAATGAGCTCCAGTCCTCTCTCGCAGAGGAGCTGGTATCCGAG
TCTTGTGAGGCCGACCCTGCGGAGCAGCTGGCCTTGCAGTCCACGCTCACTGTCTTAGCCGAGCGAATGT
CCACCATCAGGATGAAAGCCTCGGGGAAACGGCAGCTTTTGGAGGAGAAGTTGAATGATCAGCTGGAGGA
ACAAAGGCAGGAACAGGCCCTGCAGAGGTATCGCTGTGAAGCCGATGAGCTGGACAGCTGGCTCTTGAGT
ACCAAGGCCACTCTGGACACTGCGCTGAGTCCACCCAAGGAGCCCATGGACATGGAGGCCCAGCTTATGG
ACTGCCAGAATATGCTGGTGGAAATAGAGCAGAAGGTGGTGGCTTTATCAGAACTGTCAGTCCACAATGA
GAACCTGCTGCTGGAGGGCAAAGCTCACACCAAGGACGAGGCCGAGCAGCTGGCTGGAAAGCTGAGAAGG
CTCAAGGGGAGCCTGCTGGAGCTGCAGAGAGCCCTGCATGATAAGCAGCTCAACATGCAGGGAACAGCAC
AGGAGAAGGAGGAGAGCGATGTTGACCTAACAGCCACGCAGAGCCCCGGCGTCCAGGAATGGCTGGCCCA
AGCTCGCACCACATGGACCCAGCAGCGGCAGAGCAGTCTCCAGCAACAAAAAGAGTTAGAACAGGAATTA
GCCGAGCAGAAGAGTCTCCTTCGCTCAGTAGCCAGTCGTGGAGAGGAGATTCTAATTCAACATTCGGCGG
CAGAGACCTCTGGTGATGCTGGCGAAAAACCTGATGTGTTATCCCAGGAGTTGGGGATGGAAGGGGAGAA
ATCATCCGCTGAAGACCAGATGAGAATGAAATGGGAAAGCCTACATCAAGAATTTAGTACCAAGCAGAAA
CTACTACAGAATGTTCTGGAACAGGAACAAGAGCAAGTGCTTTATAGCAGGCCAAATCGACTCTTGTCTG
GTGTGCCACTGTACAAAGGGGACGTGCCAACCCAAGATAAATCTGCAGTTACATCTTTGCTGGATGGACT
GAACCAAGCCTTCGAGGAGGTTTCATCCCAGAGTGGAGGGGCAAAGAGGCAGAGTATACACTTGGAGCAG
AAGTTGTATGATGGAGTCTCAGCCACCTCTACTTGGTTGGATGACGTTGAAGAACGTTTATTTGTTGCCA
CAGCACTTTTACCAGAAGAAACAGAGACTTGTCTCTTCAACCAAGAGATTCTTGCCAAAGACATTAAGGA
AATGTCTGAAGAAATGGATAAGAACAAAAACTTGTTTTCCCAAGCTTTTCCAGAGAATGGTGATAATCGA
GATGTTATTGAAGATACTTTGGGTTGTCTTTTGGGCAGGTTATCCTTGCTAGACTCAGTAGTGAATCAAC

GATGTCATCAGATGAAAGAAAGACTTCAGCAAATACTAAATTTCCAGAATGATCTGAAAGTGCTGTTTAC
ATCACTGGCTGACAACAAATACATCATTCTGCAAAAACTGGCAAATGTGTTTGAACAGCCCGTAGCAGAA
CAAATAGAGGCAATACAACAGGCTGAAGATGGACTCAAAGAATTTGATGCAGGAATCATTGAATTAAAGA
GGCGTGGTGACGAGCTACAGGTCGAGCAGCCGTCCATGCAAGAACTCTCCAAGCTCCAGGACATGTATGA
TGAGCTGATGATGATCATTGGCTCCCGGAGGAGTGGTCTGAATCAGAACCTTACACTCAAGAGTCAGTAT
GAGAGGGCCCTACAAGATCTGGCTGACCTGCTAGAAACTGGTCAGGAGAAGATGGCAGGAGACCAGAAAA
TCATCGTGTCTTCCAAAGAGGAAATCCAGCAACCACTTGACAAACATAAGGAATACTTTCAGGGCCTGGA
ATCTCATATGATCTTGACTGTAACACTCTTCAGAAAGATAATCAGCTTTGCAGTCCAAAAGGAAACCCAG
TTCCATACAGAGCTGATGGCTCAGGCTTCTGCTGTACTGAAACGGGCTCACAAGAGGGGTGTGGAGCTGG
AGTACATTCTAGAGACGTGGTCCCATCTGGATGAGGACCAGCAGGAGCTCAGCAGACAGCTGGAGGTGGT
GGAAAGCAGCATCCCAAGCGTGGGTCTGGTGGAGGAGAACGAGGACAGGCTTATTGACCGCATAACACTC
TACCAGCATTTAAAATCTAGCCTTAATGAATACCAGCCCAAATTATATCAAGTATTAGATGATGGGAAAC
GACTTCTGATATCCATCAGCTGCTCAGATCTAGAAAGCCAACTAAATCAACTTGGAGAGTGCTGGCTAAG
TAACACCAATAAAATGTCTAAGGAACTTCACAGACTGGAAACAATATTGAAACACTGGACCAGATATCAA
AGTGAATCTGCAGATCTAATTCACTGGTTACAATCTGCAAAAGACCGGCTAGAATTTTGGACTCAGCAAT
CTGTGACAGTCCCACAAGAGCTGGAAATGGTCCGTGATCATCTAAATGCTTTCCTGGAGTTTTCTAAAGA
AGTGGATGCCCAATCTTCCCTGAAATCATCTGTTCTGAGTACTGGAAATCAGCTCCTTCGACTAAAAAAG
GTGGACACAGCCACGCTGCGCTCTGAGTTGTCGCGCATTGATAGCCAGTGGACTGACCTGCTAACCAATA
TCCCAGCCGTCCAGGAGAAGCTCCACCAGCTCCAGATGGATAAACTGCCTTCCCGCCATGCCATTTCTGA
AGTCATGAGTTGGACTTCTCTAATGGAAAATGCTATTCAGAAGGATGAAGATAATATTAAAAATTCCATA
GGTTACAAGGCAATTCATGAATACCTTCAGAAATATAAGGGTTTTAAGATAGACATTAACTGTAAACAGC
TGACAGTGGATTTTGTGAACCAGTCCGTGCTACAAATCAGCAGTCAGGATGTGGAAAGTAAGCGTAGTGA
TAAGACTGATTTTGCTGAGCAACTTGGAGCAATGAATAAAAGTTGGCAAATTCTGCAAGGTCTAGTAACT
GAGAAGATCCAGCTGTTGGAAGGCTTATTGGAATCTTGGTCAGAATATGAAAATAATGTACAATGTCTGA
AAACATGGTTTGAAACCCAGGAAAAGAGACTAAAACAACAGCATCGAATTGGAGATCAGGCTTCTGTTCA
AAATGCACTGAAAGACTGTCAGGATCTGGAAGATCTGATTAAAGCAAAAGATAAAGAAGTAGAGAAAATT
GAGCAGAATGGACTTGCTTTGATTCAGACCAAGAAAGAAGACGTCTCTAGCATTGTCATGAGCACACTGC
GAGAGCTCGGCCAAACCTGGGCAAATTTAGATCACATGGTTGGACAATTAAAGATACTGCTGAAATCAGT
GCTTGACCAATGGAGTAGTCACAAAGTGGCCTTTGACAAGATAAACAGTTACCTCATGGAGGCCAGATAC
TCTCTTTCCCGATTCCGTCTGCTGACTGGCTCCTTAGAAGCTGTGCAAGTTCAGGTGGACAATCTTCAGA
ATCTCCAAGATGATCTGGAAAAACAGGAAAGGAGCTTACAGAAATTTGGCTCTATCACCAACCAATTATT
AAAAGAGTGTCACCCACCCGTGACAGAAACTCTTACCAATACACTGAAAGAAGTCAACATGAGATGGAAT
AACTTGCTGGAAGAGATTGCTGAGCAGCTACAGTCCAGCAAGGCCCTACTTCAGCTTTGGCAAAGATACA
AGGACTACTCCAAACAGTGTGCTTCGACAGTTCAGCAGCAGGAGGATCGAACCAATGAGCTGTTGAAGGC
AGCCACAAACAAGGACATTGCCGATGATGAGGTTGCCACATGGATTCAAGATTGCAACGACCTCCTCAAA
GGACTGGGCACAGTTAAAGATTCCCTCTTTGTTCTCCATGAGCTGGGAGAGCAACTGAAGCAACAAGTGG
ATGCTTCCGCAGCATCAGCTATTCAATCGGATCAACTCTCTTTGAGTCAACACTTGTGTGCCCTGGAGCA
AGCTCTCTGCAAACAGCAGACTTCATTACAGGCTGGAGTTCTTGATTATGAAACCTTTGCCAAGAGTTTA
GAAGCTTTGGAGGCCTGGATAGTGGAAGCTGAAGAAATACTACAAGGGCAGGACCCTAGCCACTCATCTG
ACCTCTCCACAATCCAGGAAAGGATGGAAGAACTTAAGGGACAGATGTTAAAATTCAGCAGCATGGCTCC
AGATTTAGACCGTCTAAATGAGCTTGGATATAGGTTACCCTTGAATGATAAGGAAATCAAAAGAATGCAG
AATCTGAACCGCCATTGGTCTCTGATCTCCTCTCAGACTACAGAAAGATTCAGCAAGTTGCAGTCATTTT
TGCTACAACATCAGACTTTCTTGGAAAAAATGTGAAACATGGATGGAATTCCTAGTTCAGACAGAACAAAA
GTTAGCAGTAGAGATTTCAGGAAATTATCAGCACCTTTTGGAACAGCAGAGAGCACACGAGTTGTTTCAA
GCCGAGATGTTCAGTCGTCAGCAGATTTTGCACTCAATCATTATTGATGGGCAACGTCTTCTAGAACAAG
GTCAAGTTGATGACAGGGATGAATTCAACCTGAAATTGACACTCCTCAGTAATCAATGGCAGGGAGTGAT
TCGCAGGGCCCAGCAGAGGCGGGGGATCATTGACAGCCAGATTCGCCAGTGGCAGCGCTATAGGGAGATG
GCAGAAAAGCTTCGTAAATGGTTGGTTGAAGTGTCCTACCTCCCCATGAGTGGTCTCGGAAGTGTTCCTA
TACCACTGCAACAAGCAAGGACCCTCTTTGATGAAGTGCAGTTCAAAGAAAAAGTGTTTCTGCGGCAACA
AGGCAGCTACATCCTGACTGTGGAGGCTGGCAAGCAACTCCTTCTCTCGGCGGACAGTGGCGCTGAGGCC
GCCTTGCAGGCCGAACTCGCTGAAATCCAAGAGAAATGGAAATCAGCCAGCATGCGGCTGGAAGAACAGA
AGAAAAAACTAGCCTTCTTGTTGAAAGACTGGGAAAAATGTGAGAAAGGAATAGCAGATTCCCTGGAGAA
ACTACGAACTTTCAAAAAGAAGCTTTCGCAGTCTCTCCCGGATCACCATGAAGAGCTCCATGCAGAACAA
ATGCGTTGCAAGGAATTAGAAAATGCAGTTGGGAGCTGGACAGATGACTTGACCCAGTTGAGCCTGCTGA
AGGACACCCTCTCTGCCTATATCAGTGCTGATGATATCTCCATTCTTAATGAACGCGTAGAGCTTCTGCA
AAGGCAGTGGGAAGAACTATGCCACCAGCTCTCCTTAAGGCGGCAGCAAATAGGTGAAAGATTGAATGAA
TGGGCAGTCTTCAGTGAAAAGAACAAGGAACTCTGTGAGTGGTTGACTCAAATGGAAAGCAAAGTTTCTC
AGAATGGAGACATTCTCATTGAAGAAATGATAGAGAAGCTCAAGAAGGATTATCAAGAGGAAATTGCTAT
TGCTCAAGAGAACAAAATACAGCTCCAACAAATGGGAGAACGACTTGCTAAAGCCAGCCATGAAAGCAAA
GCATCTGAGATTGAATACAAGCTGGGAAAGGTCAACGACCGGTGGCAGCATCTCCTGGACCTCATTGCAG
CCAGGGTGAAGAAGCTGAAGGAGACCCTGGTAGCCGTGCAGCAGCTTGATAAGAACATGAGCAGCCTGAG
GACCTGGCTCGCTCACATCGAGTCAGAGCTGGCCAAGCCAATAGTCTACGATTCCTGTAACTCGGAAGAA
ATACAGAGAAAGCTTAATGAGCAGCAGGAGCTTCAGAGAGACATAGAGAAGCACAGTACAGGTGTTGCAT

CTGTCCTCAACCTGTGTGAAGTCCTGCTGCACGACTGTGACGCCTGTGCCACTGATGCCGAGTGTGACTC
TATACAGCAGGCTACGAGAAACCTGGACCGGCGGTGGAGAAACATTTGTGCTATGTCCATGGAAAGGAGG
CTGAAAATCGAAGAGACGTGGCGATTGTGGCAGAAATTTCTGGATGACTATTCACGTTTTGAAGATTGGC
TGAAGTCTTCAGAAAGGACAGCTGCTTTTCCCAGCTCTTCTGGGGTGATCTATACAGTTGCCAAGGAAGA
ACTAAAGAAATTTGAGGCTTTCCAGCGACAGGTCCACGAGTGCCTGACGCAGCTGGAACTGATCAACAAG
CAGTACCGCCGCCTGGCCAGGGAGAACCGCACTGATTCAGCATGTAGCCTCAAACAGATGGTTCACGAAG
GCAACCAGAGATGGGACAACCTGCAAAAGCGTGTCACCTCCATCTTGCGCAGACTCAAGCATTTTATTGG
CCAGCGTGAGGAGTTTGAGACTGCGCGGGACAGCATTCTGGTCTGGCTCACAGAGATGGATCTGCAGCTC
ACTAATATTGAACATTTTTCTGAGTGTGATGTTCAAGCTAAAATAAAGCAACTCAAGGCCTTCCAGCAGG
AAATTTCACTGAACCACAATAAGATTGAGCAGATAATTGCCCAAGGAGAACAGCTGATAGAAAAGAGTGA
GCCCTTGGATGCAGCGATCATCGAGGAGGAACTAGATGAGCTCCGACGGTACTGCCAGGAGGTCTTCGGG
CGTGTGGAAAGATACCATAAGAAACTGATCCGCCTGCCTCTCCCAGACGATGAGCACGACCTCTCAGACA
GGGAGCTGGAGCTGGAAGACTCTGCAGCTCTGTCGGACCTGCACTGGCACGACCGCTCTGCAGACAGCCT
GCTTTCTCCACAGCCTTCCTCCAATCTCTCCCTCTCGCTCGCTCAGCCCCTCCGGAGCGAGCGGTCAGGA
CGAGACACCCCAGCTAGTGTGGACTCCATCCCCCTGGAGTGGGATCACGACTATGACCTCAGTCGGGACC
TGGAGTCTGCAATGTCCAGAGCTCTGCCCTCTGAGGATGAAGAAGGTCAGGATGACAAAGATTTCTACCT
CCGGGGAGCTGTTGCCTTATCAGGGGACCACAGTGCCCTAGAGTCACAGATCCGACAACTGGGCAAAGCC
CTGGATGATAGCCGCTTTCAGATACAGCAAACCGAAAATATCATTCGCAGCAAAACTCCCACGGGGCCGG
AGCTAGACACCAGCTACAAAGGCTACATGAAACTGCTGGGCGAATGCAGTAGCAGTATAGACTCCGTGAA
GAGACTGGAGCACAAACTGAAGGAGGAAGAGGAGAGCCTTCCTGGCTTTGTTAACCTGCATAGTACCGAA
ACCCAAACGGCTGGTGTGATTGACCGATGGGAGCTTCTCCAGGCCCAGGCATTGAGCAAGGAGTTGAGGA
TGAAGCAGAACCTCCAGAAGTGGCAGCAGTTTAACTCAGACTTGAACAGCATCTGGGCCTGGCTGGGGGA
CACGGAGGAGGAGTTGGAACAGCTCCAGCGTCTGGAACTCAGCACTGACATCCAGACCATCGAGCTCCAG
ATCAAAAAGCTCAAGGAGCTCCAGAAAGCTGTGGACCACCGCAAAGCCATCATCCTCTCCATCAATCTCT
GCAGCCCTGAGTTCACCCAGGCTGACAGCAAGGAGAGCCGGGACCTGCAGGATCGCTTGTCGCAGATGAA
TGGGCGCTGGGACCGAGTGTGCTCTCTGCTGGAGGAGTGGCGGGGCCTGCTGCAGGATGCCCTGATGCAG
TGCCAGGGTTTCCATGAAATGAGCCATGGTTTGCTTCTTATGCTGGAGAACATTGACAGAAGGAAAAATG
AAATTGTCCCTATTGATTCTAACCTTGATGCAGAGATACTTCAGGACCATCACAAACAGCTTATGCAAAT
AAAGCATGAGCTGTTGGAATCCCAACTCAGAGTAGCCTCTTTGCAAGACATGTCTTGCCAACTACTGGTG
AATGCTGAAGGAACAGACTGTTTAGAAGCCAAAGAAAAAGTCCATGTTATTGGAAATCGGCTCAAACTTC
TCTTGAAGGAGGTCAGTCGTCATATCAAGGAACTGGAGAAGTTATTAGACGTGTCAAGTAGTCAGCAGGA
TTTGTCTTCCTGGTCTTCTGCTGATGAACTGGACACCTCAGGGTCTGTGAGTCCCACATCAGGAAGGAGC
ACCCCAAACAGACAGAAAACGCCACGAGGCAAGTGTAGTCTCTCACAGCCTGGACCCTCTGTCAGCAGTC
CACATAGCAGGTCCACAAAAGGTGGCTCCGATTCCTCCCTTTCTGAGCCAGGGCCAGGTCGGTCCGGCCG
CGGCTTCCTGTTCAGAGTCCTCCGAGCAGCTCTTCCCCTTCAGCTTCTCCTGCTCCTCCTCATCGGGCTT
GCCTGCCTTGTACCAATGTCAGAGGAAGACTACAGCTGTGCCCTCTCCAACAACTTTGCCCGGTCATTCC
ACCCCATGCTCAGATACACGAATGGCCCTCCTCCACTCTGAACTAAGCAGATGCCATCTGCAGAAGTGCT
GGTAGCATAAGGAGGATCGGGTCATAAGCAATCCCAAACTACCAACAAGAGGACCTTGATCTTGGCGAAA
GCCCTCGGTGTGGCAGCTTTAGCCCTCCTCCAGATCACATGTGTGCAAATTATGGCTTCAGAGGTGGAAG
ATAAACAGTGACGGGGGAACAAACAGACAACAAGAAGGTTTGGAAGAAATCTGGTTTGAGACTCTGAACC
TTAGCACTAAGGAGATTGAGTAAGGACCTCCAAAGTTCCCCGGACTCATGAATTCTGGGCCCTTGGCCCA
TTCTGTGCACAGCCAAGGACTTCAGTAGACCATCTGGGCAGCTTTCCCATGGTGCTGCTCCAACCATCAG
ATAAATGACCCTCCCAAGCACCATGTCAGTGTCGTACAATCTACCAACCAACCAGTGCTGAAGAGATTTT
AGAACCTTGTAACATACAATTTTTAAGAGCTTATATGGCAGCTTCCTTTTTACCTTGTTTTCCTTTGGGG
CATGATGTTTTAACCTTTGCTTTAGAAGCACAAGCTGTAAATCTAAAAGGCACTTTTTTTTAGAGGTATA
AAGAAAAACTAGATGTAATAAATAAGATCATGGAAGGCTTTATGTGAAAAAAGTTGAATGTTATAGT

Human SYNE1 mRNA sequence - var19 (public gi: 28195676) (SEQ ID NO: 201)
CAAGGGGAAACTTTCATCCCCACGCAGGTTATAGCTTTTTGCTCTGCAGAGTCTAACTTTTGCAAGTGGA
AGCTTCATGGTGGTGGCGGAGGACCTGAGTGCCCTGAGGATGGCAGAGGACGGCTGTGTGGATGCAGATC
TCCCAGATTGTAACTGCGATGTCACAAGGGCCAGGGTGAAGAAGCTGAAGGAGACCCTGGTAGCCGTGCA
GCAGCTTGATAAGAACATGAGCAGCCTGAGGACCTGGCTCGCTCACATCGAGTCAGAGCTGGCCAAGCCA
ATAGTCTACGATTCCTGTAACTCGGAAGAAATACAGAGAAAGCTTAATGAGCAGCAGGAGCTTCAGAGAG
ACATAGAGAAGCACAGTACAGGTGTTGCATCTGTCCTCAACCTGTGTGAAGTCCTGCTGCACGACTGTGA
CGCCTGTGCCACTGATGCCGAGTGTGACTCTATACAGCAGGCTACGAGAAACCTGGACCGGCGGTGGAGA
AACATTTGTGCTATGTCCATGGAAAGGAGGCTGAAAATCGAAGAGACGTGGCGATTGTGGCAGAAATTTC
TGGATGACTATTCACGTTTTGAAGATTGGCTGAAGTCTTCAGAAAGGACAGCTGCTTTTCCCAGCTCTTC
TGGGGTGATCTATACAGTTGCCAAGGAAGAACTAAAGAAATTTGAGGCTTTCCAGCGACAGGTCCACGAG
TGCCTGACGCAGCTGGAACTGATCAACAAGCAGTACCGCCGCCTGGCCAGGGAGAACCGCACTGATTCAG
CATGTAGCCTCAAACAGATGGTTCACGAAGGCAACCAGAGATGGGACAACCTGCAAAAGCGTGTCACCTC
CATCTTGCGCAGACTCAAGCATTTTATTGGCCAGCGTGAGGAGTTTGAGACTGCGCGGGACAGCATTCTG
GTCTGGCTCACAGAGATGGATCTGCAGCTCACTAATATTGAACATTTTTCTGAGTGTGATGTTCAAGCTA
AAATAAAGCAACTCAAGGCCTTCCAGCAGGAAATTTCACTGAACCACAATAAGATTGAGCAGATAATTGC

CCAAGGAGAACAGCTGATAGAAAAGAGTGAGCCCTTGGATGCAGCGATCATCGAGGAGGAACTAGATGAG
CTCCGACGGTACTGCCAGGAGGTCTTCGGGCGTGTGGAAAGATACCATAAGAAACTGATCCGCCTGCCTC
TCCCAGACGATGAGCACGACCTCTCAGACAGGGAGCTGGAGCTGGAAGACTCTGCAGCTCTGTCGGACCT
GCACTGGCACGACCGCTCTGCAGACAGCCTGCTTTCTCCACAGCCTTCCTCCAATCTCTCCCTCTCGCTC
GCTCAGCCCCTCCGGAGCGAGCGGTCAGGACGAGACACCCCAGCTAGTGTGGACTCCATCCCCCTGGAGT
GGGATCACGACTATGACCTCAGTCGGGACCTGGAGTCTGCAATGTCCAGAGCTCTGCCCTCTGAGGATGA
AGAAGGTCAGGATGACAAAGATTTCTACCTCCGGGGAGCTGTTGCCTTATCAGATGTAATGATCCCCGAA
AGCCCTGAGGCCTATGTAAAACTCACAGAAAATGCAATCAAAAATACCTCCGGGGACCACAGTGCCCTAG
AGTCACAGATCCGACAACTGGGCAAAGCCCTGGATGATAGCCGCTTTCAGATACAGCAAACCGAAAATAT
CATTCGCAGCAAAACTCCCACGGGGCCGGAGCTAGACACCAGCTACAAAGGCTACATGAAACTGCTGGGC
GAATGCAGTAGCAGTATAGACTCCGTGAAGAGACTGGAGCACAAACTGAAGGAGGAAGAGGAGAGCCTTC
CTGGCTTTGTTAACCTGCATAGTACCGAAACCCAAACGGCTGGTGTGATTGACCGATGGGAGCTTCTCCA
GGCCCAGGCATTGAGCAAGGAGTTGAGGATGAAGCAGAACCTCCAGAAGTGGCAGCAGTTTAACTCAGAC
TTGAACAGCATCTGGGCCTGGCTGGGGACACGGAGGAGGAGTTGGAACAGCTCCAGCGTCTGGAACTCA
GCACTGACATCCAGACCATCGAGCTCCAGATCAAAAAGCTCAAGGAGCTCCAGAAAGCTGTGGACCACCG
CAAAGCCATCATCCTCTCCATCAATCTCTGCAGCCCTGAGTTCACCCAGGCTGACAGCAAGGAGAGCCGG
GACCTGCAGGATCGCTTGTCGCAGATGAATGGGCGCTGGGACCGAGTGTGCTCTCTGCTGGAGGAGTGGC
GGGGCCTGCTGCAGGATGCCCTGATGCAGTGCCAGGGTTTCCATGAAATGAGCCATGGTTTGCTTCTTAT
GCTGGAGAACATTGACAGAAGGAAAAATGAAATTGTCCCTATTGATTCTAACCTTGATGCAGAGATACTT
CAGGACCATCACAAACAGCTTATGCAAATAAAGCATGAGCTGTTGGAATCCCAACTCAGAGTAGCCTCTT
TGCAAGACATGTCTTGCCAACTACTGGTGAATGCTGAAGGAACAGACTGTTTAGAAGCCAAAGAAAAAGT
CCATGTTATTGGAAATCGGCTCAAACTTCTCTTGAAGGAGGTCAGTCGTCATATCAAGGAACTGGAGAAG
TTATTAGACGTGTCAAGTAGTCAGCAGGATTTGTCTTCCTGGTCTTCTGCTGATGAACTGGACACCTCAG
GGTCTGTGAGTCCCACATCAGGAAGGAGCACCCCAAACAGACAGAAAACGCCACGAGGCAAGTGTAGTCT
CTCACAGCCTGGACCCTCTGTCAGCAGTCCACATAGCAGGTCCACAAAAGGTGGCTCCGATTCCTCCCTT
TCTGAGCCAGGGCCAGGTCGGTCCGGCCGCGGCTTCCTGTTCAGAGTCCTCCGAGCAGCTCTTCCCCTTC
AGCTTCTCCTGCTCCTCCTCATCGGGCTTGCCTGCCTTGTACCAATGTCAGAGGAAGACTACAGCTGTGC
CCTCTCCAACAACTTTGCCCGGTCATTCCACCCCATGCTCAGATACACGAATGGCCCTCCTCCACTCTGA
ACTAAGCAGATGCCATCTGCAGAAGTGCTGGTAGCATAAGGAGGATCGGGTCATAAGCAATCCCAAACTA
CCAACAAGAGGACCTTGATCTTGGCGAAAGCCCTCGGTGTGGCAGCTTTAGCCCTCCTCCAGATCACATG
TGTGCAAATTATGGCTTCAGAGGTGGAAGATAAACAGTGACGGGGAACAAACAGACAACAAGAAGGTTT
GGAAGAAATCTGGTTTGAGACTCTGAACCTTAGCACTAAGGAGATTGAGTAAGGACCTCCAAAGTTCCCC
GGACTCATGAATTCTGGGCCCTTGGCCCATTCTGTGCACAGCCAAGGACTTCAGTAGACCATCTGGGCAG
CTTTCCCATGGTGCTGCTCCAACCATCAGATAAATGACCCTCCCAAGCACCATGTCAGTGTCGTACAATC
TACCAACCAACCAGTGCTGAAGAGATTTTAGAACCTTGTAACATACAATTTTTAAGAGCTTATATGGCAG
CTTCCTTTTTACCTTGTTTTCCTTTGGGGCATGATGTTTTAACCTTTGCTTTAGAAGCACAAGCTGTAAA
TCTAAAAGGCACTTTTTTTTAGAGGTATAAAGAAAAACTAGATGTAATAAATAAGATCATGGAAGGCTTT
ATGTGAAAAAAGTTGAATGTTATAGT

Human SYNE1 Protein sequence - var1 (public gi: 21753085) (SEQ ID NO: 295)
MVVDDLFEDMKDGVKLLALLEVLSGQKLPCEQGRRMKRIHAVANIGTALKFLEGRKIKLVNINSTDIADG
RPSIVLGLMWTIILYFQIEELTSNLPQLQSLSSSASSVDSIVSSETPSPPSKRKVTTKIQGNAKKALLKW
VQYTAGKQTGIEVKDFGKSWRSGVAFHSVIHAIRPELVDLETVKGRSNRENLEDAFTIAETELGIPRLLD
PEDVDVDKPDEKSIMTYVAQFLKHYPDIHNASTDGQEDDEILPGFPSFANSVQNFKREDRVIFKEMKVWI
EQFERDLTRAQMVESNLQDKYQSFKHFRVQYEMKRKQIEHLIQPLHRDGKLSLDQALVKQSWDRVTSRLF
DWHIQLDKSLPAPLGTIGAWLYRAEVALREEITVQQVHEETANTIQRKLEQHKDLLQNTDAHKRAFHEIY
RTRSVNGIPVPPDQLEDMAERFHFVSSTSELHLMKMEFLELKYRLLSLLVLAESKLKSWIIKYGRRESVE
QLLQNYVSFIENSKFFEQYEVTYQILKQTAEMYVKADGSVEEAENVMKFMNETTAQWRNLSVEVRSVRSM
LEEVISNWDRYGNTVASLQAWLEDAEKMLNQSENAKKDFFRNLPHWIQQHTAMNDAGNFLIETCDEMVSR
DLKQQLLLLNGRWRELFMEVKQYAQADEMDRMKKEYTDCVVTLSAFATEAHKKLSEPLEVSFMNVKLLIQ
DLEDIEQRVPVMDAQYKIITKTAHLITKESP

Human SYNE1 Protein sequence - var2 (public gi: 19584385) (SEQ ID NO: 296)
KLLIQDLEDIEQRVPVMDAQYKIITKTAHLITKESPQEEGKEMFATMSKLKEQLTKVKECYSPLLYESQQ
LLIPLEELEKQMTSFYDSLGKINEIITVLEREAQSSALFKQKHQELLACQENCKKTLTLIEKGSQSVQKF
VTLSNVLKHFDQTRLQRQIADIHVAFQSMVKKTGDWKKHVETNSRLMKKFEESRAELEKVLRIAQEGLEE
KGDPEELLRRHTEFFSQLDQRVLNAFLKACDELTDILPEQEQQGLQEAVRKLHKQWKDLQGEAPYHLLHL
KIDVEKNRFLASVEECRTELDRETKLMPQEGSEKIIKEHRVFFSDKGPHHLCEKRLQLIEELCVKLPVRD
PVRDTPGTCHVTLKELRAAIDSTYRKLMEDPDKWKDYTSRFSEFSSWISTNETQLKGIKGEAIDTANHGE
VKRAVEEIRNGVTKRGETLSWLKSRLKVLTEVSSENEAQKQGDELAKLSSSFKALVTLLSEVEKMLSNFG
DCVQYKEIVKNSLEELISGSKEVQEQAEKILDTENLFEAQQLLLHHQQKTKRISAKKRDVQQQIAQAQQG
EGGLPDRGHEELRKLESTLDGLERSRERQERRIQVTLRKWERFETNKETVVRYLFQTGSSHERFLSFSSL

ESLSSELEQTKEFSKRTESIAVQAENLVKEASEIPLGPQNKQLLQQQAKSIKEQVKKLEDTLEEDIKPME
MVKTKWDHFGSNFETLSVWITEKEKELNALETSSSAMDMQISQIKVTIQEIESKLSSIVGLEEEAQSFAQ
FVTTGESARIKAKLTQIRRYGEELREHAQCLEGTILGHLSQQQKFEENLRKIQQSVSEFEDKLAVPIKIC
SSATETYKVLQEHMDLCQALESLSSAITAFSASARKVVNRDSCVQEAAALQQQYEDILRRAKERQTALEN
LLAHWQRLEKELSSFLTWLERGEAKASSPEMDISADRVKVEGELQLIQASSRKCEEGKNKMLFVTVTLFK
IIK

Human SYNE1 Protein sequence - var3 (public gi: 17861378) (SEQ ID NO: 297)
MGERLAKASHESKASEIEYKLGKVNDRWQHLLDLIAARVKKLKETLVAVQQLDKNMSSLRTWLAHIESEL
AKPIVYDSCNSEEIQRKLNEQQELQRDIEKHSTGVASVLNLCEVLLHDCDACATDAECDSIQQATRNLDR
RWRNICAMSMERRLKIEETWRLWQKFLDDYSRFEDWLKSSERTAAFPSSSGVIYTVAKEELKKFEAFQRQ
VHECLTQLELINKQYRRLARENRTDSACSLKQMVHEGNQRWDNLQKRVTSILRRLKHFIGQREEFETARD
SILVWLTEMDLQLTNIEHFSECDVQAKIKQLKAFQQEISLNHNKIEQIIAQGEQLIEKSEPLDAAIIEEE
LDELRRYCQEVFGRVERYHKKLIRLPLPDDEHDLSDRELELEDSAALSDLHWHDRSADSLLSPQPSSNLS
LSLAQPLRSERSGRDTPASVDSIPLEWDHDYDLSRDLESAMSRALPSEDEEGQDDKDFYLRGAVALSDVM
IPESPEAYVKLTENAIKNTSGDHSALESQIRQLGKALDDSRFQIQQTENIIRSKTPTGPELDTSYKGYMK
LLGECSSSIDSVKRLEHKLKEEEESLPGFVNLHSTETQTAGVIDRWELLQAQALSKELRMKQNLQKWQQF
NSDLNSIWAWLGDTEEELEQLQRLELSTDIQTIELQIKKLKELQKAVDHRKAIILSINLCSPEFTQADSK
ESRDLQDRLSQMNGRWDRVCSLLEEWRGLLQDALMQCQGFHEMSHGLLLMLENIDRRKNEIVPIDSNLDA
EILQDHHKQLMQIKHELLESQLRVASLQDMSCQLLVNAEGTDCLEAKEKVHVIGNRLKLLLKEVSRHIKE
LEKLLDVSSSQQDLSSWSSADELDTSGSVSPTSGRSTPNRQKTPRGKCSLSQPGPSVSSPHSRSTKGGSD
SSLSEPGPGRSGRGFLFRVLRAALPLQLLLLLLIGLACLVPMSEEDYSCALSNNFARSFHPMLRYTNGPP
PL

Human SYNE1 Protein sequence - var4 (public gi: 17861386) (SEQ ID NO: 298)
MELDAAVQKFLEQNGQLGKPLAKKIGKLTELHQQTIRQAENRLSKLNQATSHLEEYNEMLELILKWIEKA
KVLAHGTIAWNSASQLRKQYILHQTLLEESKEIDSELEAMTEKLQYLTSVYCTEKMSQQVAELGRETEEL
RQMIKIRLQNLQDAAKDMKKFEAELKKLQAALEQAQATLTSPEVGRLSLKEQLSHRQHLLSEMESLKPKV
QAVQLCQSALRIPEDVVASLPLCHAALRLQEEASRLQHTAIQQCNIMQEAVVQYEQYEQEMKHLQQLIEG
AHREIEDKPVATSNIQELQAQISRHEELAQKIKGYQEQIASLNSKCKMLTMKAKHATMLLTVTEVEGLAE
GTEDLDGELLPTPSAHPSVVMMTAGRCHTLLSPVTEESGEEGTNSEISSPPACRSPSPVANTDASVNQDI
AYYQALSAERLQTDAAKIHPSTSASQEFYEPGLEPSATAKLGDLQRSWETLKNVISEKQRTLYEALERQQ
KYQDSLQSISTKMEAIELKLSESPEPGRSPESQMAEHQALMDEILMLQDEINELQSSLAEELVSESCEAD
PAEQLALQSTLTVLAERMSTIRMKASGKRQLLEEKLNDQLEEQRQEQALQRYRCEADELDSWLLSTKATL
DTALSPPKEPMDMEAQLMDCQNMLVEIEQKVVALSELSVHNENLLLEGKAHTKDEAEQLAGKLRRLKGSL
LELQRALHDKQLNMQGTAQEKEESDVDLTATQSPGVQEWLAQARTTWTQQRQSSLQQQKELEQELAEQKS
LLRSVASRGEEILIQHSAAETSGDAGEKPDVLSQELGMEGEKSSAEDQMRMKWESLHQEFSTKQKLLQNV
LEQEQEQVLYSRPNRLLSGVPLYKGDVPTQDKSAVTSLLDGLNQAFEEVSSQSGGAKRQSIHLEQKLYDG
VSATSTWLDDVEERLFVATALLPEETETCLFNQEILAKDIKEMSEEMDKNKNLFSQAFPENGDNRDVIED
TLGCLLGRLSLLDSVVNQRCHQMKERLQQILNFQNDLKVLFTSLADNKYIILQKLANVFEQPVAEQIEAI
QQAEDGLKEFDAGIIELKRRGDELQVEQPSMQELSKLQDMYDELMMIIGSRRSGLNQNLTLKSQYERALQ
DLADLLETGQEKMAGDQKIIVSSKEEIQQPLDKHKEYFQGLESHMILTVTLFRKIISFAVQKETQFHTEL
MAQASAVLKRAHKRGVELEYILETWSHLDEDQQELSRQLEVVESSIPSVGLVEENEDRLIDRITLYQHLK
SSLNEYQPKLYQVLDDGKRLLISISCSDLESQLNQLGECWLSNTNKMSKELHRLETILKHWTRYQSESAD
LIHWLQSAKDRLEFWTQQSVTVPQELEMVRDHLNAFLEFSKEVDAQSSLKSSVLSTGNQLLRLKKVDTAT
LRSELSRIDSQWTDLLTNIPAVQEKLHQLQMDKLPSRHAISEVMSWTSLMENAIQKDEDNIKNSIGYKAI
HEYLQKYKGFKIDINCKQLTVDFVNQSVLQISSQDVESKRSDKTDFAEQLGAMNKSWQILQGLVTEKIQL
LEGLLESWSEYENNVQCLKTWFETQEKRLKQQHRIGDQASVQNALKDCQDLEDLIKAKDKEVEKIEQNGL
ALIQTKKEDVSSIVMSTLRELGQTWANLDHMVGQLKILLKSVLDQWSSHKVAFDKINSYLMEARYSLSRF
RLLTGSLEAVQVQVDNLQNLQDDLEKQERSLQKFGSITNQLLKECHPPVTETLTNTLKEVNMRWNNLLEE
IAEQLQSSKALLQLWQRYKDYSKQCASTVQQQEDRTNELLKAATNKDIADDEVATWIQDCNDLLKGLGTV
KDSLFVLHELGEQLKQQVDASAASAIQSDQLSLSQHLCALEQALCKQQTSLQAGVLDYETFAKSLEALEA
WIVEAEEILQGQDPSHSSDLSTIQERMEELKGQMLKFSSMAPDLDRLNELGYRLPLNDKEIKRMQNLNRH
WSLISSQTTERFSKLQSFLLQHQTFLEKCETWMEFLVQTEQKLAVEISGNYQHLLEQQRAHELFQAEMFS
RQQILHSIIIDGQRLLEQGQVDDRDEFNLKLTLLSNQWQGVIRRAQQRRGIIDSQIRQWQRYREMAEKLR
KWLVEVSYLPMSGLGSVPIPLQQARTLFDEVQFKEKVFLRQQGSYILTVEAGKQLLLSADSGAEAALQAE
LAEIQEKWKSASMRLEEQKKKLAFLLKDWEKCEKGIADSLEKLRTFKKKLSQSLPDHHEELHAEQMRCKE
LENAVGSWTDDLTQLSLLKDTLSAYISADDISILNERVELLQRQWEELCHQLSLRRQQIGERLNEWAVFS
EKNKELCEWLTQMESKVSQNGDILIEEMIEKLKKDYQEEIAIAQENKIQLQQMGERLAKASHESKASEIE
YKLGKVNDRWQHLLDLIAARVKKLKETLVAVQQLDKNMSSLRTWLAHIESELAKPIVYDSCNSEEIQRKL
NEQQELQRDIEKHSTGVASVLNLCEVLLHDCDACATDAECDSIQQATRNLDRRWRNICAMSMERRLKIEE
TWRLWQKFLDDYSRFEDWLKSSERTAAFPSSSGVIYTVAKEELKKFEAFQRQVHECLTQLELINKQYRRL

ARENRTDSACSLKQMVHEGNQRWDNLQKRVTSILRRLKHFIGQREEFETARDSILVWLTEMDLQLTNIEH
FSECDVQAKIKQLKAFQQEISLNHNKIEQIIAQGEQLIEKSEPLDAAIIEEELDELRRYCQEVFGRVERY
HKKLIRLPLPDDEHDLSDRELELEDSAALSDLHWHDRSADSLLSPQPSSNLSLSLAQPLRSERSGRDTPA
SVDSIPLEWDHDYDLSRDLESAMSRALPSEDEEGQDDKDFYLRGAVALSGDHSALESQIRQLGKALDDSR
FQIQQTENIIRSKTPTGPELDTSYKGYMKLLGECSSSIDSVKRLEHKLKEEEESLPGFVNLHSTETQTAG
VIDRWELLQAQALSKELRMKQNLQKWQQFNSDLNSIWAWLGDTEEELEQLQRLELSTDIQTIELQIKKLK
ELQKAVDHRKAIILSINLCSPEFTQADSKESRDLQDRLSQMNGRWDRVCSLLEEWRGLLQDALMQCQGFH
EMSHGLLLMLENIDRRKNEIVPIDSNLDAEILQDHHKQLMQIKHELLESQLRVASLQDMSCQLLVNAEGT
DCLEAKEKVHVIGNRLKLLLKEVSRHIKELEKLLDVSSSQQDLSSWSSADELDTSGSVSPTSGRSTPNRQ
KTPRGKCSLSQPGPSVSSPHSRSTKGGSDSSLSEPGPGRSGRGFLFRVLRAALPLQLLLLLLIGLACLVP
MSEEDYSCALSNNFARSFHPMLRYTNGPPPL

Human SYNE1 Protein sequence - var5 (public gi: 17227154) (SEQ ID NO: 299)
MRLEEQKKKLAFLLKDWEKCEKGIADSLEKLRTFKKKLSQSLPDHHEELHAEQMRCKELENAVGSWTDDL
TQLSLLKDTLSAYISADDISILNERVELLQRQWEELCHQLSLRRQQIGERLNEWAVFSEKNKELCEWLTQ
MESKVSQNGDILIEEMIEKLKKDYQEEIAIAQENKIQLQQMGERLAKASHESKASEIEYKLGKVNDRWQH
LLDLIAARVKKLKETLVAVQQLDKNMSSLRTWLAHIESELAKPIVYDSCNSEEIQRKLNEQQELQRDIEK
HSTGVASVLNLCEVLLHDCDACATDAECDSIQQATRNLDRRWRNICAMSMERRLKIEETWRLWQKFLDDY
SRFEDWLKSSERTAAFPSSSGVIYTVAKEELKKFEAFQRQVHECLTQLELINKQYRRLARENRTDSACSL
KQMVHEGNQRWDNLQKRVTSILRRLKHFIGQREEFETARDSILVWLTEMDLQLTNIEHFSECDVQAKIKQ
LKAFQQEISLNHNKIEQIIAQGEQLIEKSEPLDAAIIEEELDELRRYCQEVFGRVERYHKKLIRLPLPDD
EHDLSDRELELEDSAALSDLHWHDRSADSLLSPQPSSNLSLSLAQPLRSERSGRDTPASVDSIPLEWDHD
YDLSRDLESAMSRALPSEDEEGQDDKDFYLRGAVALSGDHSALESQIRQLGKALDDSRFQIQQTENIIRS
KTPTGPELDTSYKGYMKLLGECSSSIDSVKRLEHKLKEEEESLPGFVNLHSTETQTAGVIDRWELLQAQA
LSKELRMKQNLQKWQQFNSDLNSIWAWLGDTEEELEQLQRLELSTDIQTIELQIKKLKELQKAVDHRKAI
ILSINLCSPEFTQADSKESRDLQDRLSQMNGRWDRVCSLLEEWRGLLQDALMQCQGFHEMSHGLLLMLEN
IDRRKNEIVPIDSNLDAEILQDHHKQLMQIKHELLESQLRVASLQDMSCQLLVNAEGTDCLEAKEKVHVI
GNRLKLLLKEVSRHIKELEKLLDVSSSQQDLSSWSSADELDTSGSVSPTSGRSTPNRQKTPRGKCSLSQP
GPSVSSPHSRSTKGGSDSSLSEPGPGRSGRGFLFRVLRAALPLQLLLLLLIGLACLVPMSEEDYSCALSN
NFARSFHPMLRYTNGPPPL

Human SYNE1 Protein sequence - var6 (public gi: 12698057) (SEQ ID NO: 300)
QRKLEQHKDLLQNTDAHKRAFHEIYRTRSVNGIPVPPDQLEDMAERFHFVSSTSELHLMKMEFLELKYRL
LSLLVLAESKLKSWIIKYGRRESVEQLLQNYVSFIENSKFFEQYEVTYQILKQTAEMYVKADGSVEEAEN
VMKFMNETTAQWRNLSVEVRSVRSMLEEVISNWDRYGNTVASLQAWLEDAEKMLNQSENAKKDFFRNLPH
WIQQHTAMNDAGNFLIETCDEMVSRDLKQQLLLLNGRWRELFMEVKQYAQADEMDRMKKEYTDCVVTLSA
FATEAHKKLSEPLEVSFMNVKLLIQDLEDIEQRVPVMDAQYKIITKTAHLITKESPQEEGKEMFATMSKL
KEQLTKVKECYSPLLYESQQLLIPLEELEKQMTSFYDSLGKINEIITVLEREAQSSALFKQKHQELLACQ
ENCKKTLTLIEKGSQSVQKFVTLSNVLKHFDQTRLQRQIADIHVAFQSMVKKTGDWKKHVETNSRLMKKF
EESRAELEKVLRIAQEGLEEKGDPEELLRRHTEFFSQLDQRVLNAFLKACDELTDILPEQEQQGLQEAVR
KLHKQWKDLQGEAPYHLLHLKIDVEKNRFLASVEECRTELDRETKLMPQEGSEKIIKEHRVFFSDKGPHH
LCEKRLQLIEELCVKLPVRDPVRDTPGTCHVTLKELRAAIDSTYRKLMEDPDKWKDYTSRFSEFSSWIST
NETQLKGIKGEAIDTANHGEVKRAVEEIRNGVTKRGETLSWLKSRLKVLTEVSSENEAQKQGDELAKLSS
SFKALVTLLSEVEKMLSNFGDCVQYKEIVKNSLEELISGSKEVQEQAEKILDTENLFEAQQLLLHHQQKT
KRISAKKRDVQQQIAQAQQGEGGLPDRGHEELRKLESTLDGLERSRERQERRIQVTLRKWERFETNKETV
VRYLFQTGSSHERFLSFSSLESLSSELEQTKEFSKRTESIAVQAENLVKEASEIPLGPQNKQLLQQQAKS
IKEQVKKLEDTLEEEYVIDKS

Human SYNE1 Protein sequence - var7 (public gi: 2895593) (SEQ ID NO: 301
MKQNLQKWQQFNSDLNSIWAWLGDTEEELEQLQRLELSTDIQTIELQIKKLKELQKAVDHRKAIILSINL
CSPEFTQADSKESRDLQDRLSQMNGRWDRVCSLLEEWRGLLQDALMQCQGFHEMSHGLLLMLENIDRRKN
EIVPIDSNLDAEILQDHHKQLMQIKHELLESQLRVASLQDMSCQLLVNAEGTDCLEAKEKVHVIGNRLKL
LLKEVSRHIKELEKLLDVSSSQQDLSSWSSADELDTSGSVSPTSGRSTPNRQKTPRGKCSLSQPGPSVSS
PHSRSTKGGSDSSLSEPGPGRSGRGFLFRVLRAALPLQLLLLLLIGLACLVPMSEEDYSCALSNNFARSS
TPCSDTRMALLHSELSRCHLQKCW

Human SYNE1 Protein sequence - var8 (public gi: 6330957) (SEQ ID NO: 302
LDLCRQSNNLCLQREEDLQRTRDYHDCMNVVEVFLEKFTTEWDNLARSDAESTAVHLEALKKLALALQER
KYAIEDLKDQKQKMIEHLNLDDKELVKEQTSHLEQRWFQLEDLIKRKIQVSVTNLEELNVVQSRFQELME
WAEEQQPNIAEALKQSPPPDMAQNLLMDHLAICSELEAKQMLLKSLIKDADRVMADLGLNERQVIQKALS
DAQSHVNCLSDLVGQRRKYLNKALSEKTQFLMAVFQATSQIQQHERKIMFREHICLLPDDVSKQVKTCKS

AQASLKTYQNEVTGLWAQGRELMKEVTEQEKSEVLGKLQELQSVYDSVLQKCSHRLQELEKNLVSRKHFK
EDFDKACHWLKQADIVTFPEINLMNESTELHTQLAKYQNILEQSPEYENLLLTLQRTGQTILPSLNEVDH
SYLSEKLNALPRQFNVIVALAKDKFYKVQEAILARKEYASLIELTTQSLSELEAQFLRMSKVPTDLAVEE
ALSLQDGCRAILDEVAGLGEAVDELNQKKEGFRSTGQPWQPDKMLHLVTLYHRLKRQTEQRVSLLEDTTS
AYQEHEKMCQQLERQLKSVKEEQSKVNEETLPAEEKLKMYHSLAGSLQDSGIVLKRVTIHLEDLAPHLDP
LAYEKARHQIQSWQGELKLLTSAIGETVTECESRMVQSIDFQTEMSRSLDWLRRVKAELSGPVYLDLNLQ
DIQEEIRKIQIHQEEVQSSLRIMNALSHKEKEKFTKAKELISADLEHSLAELSELDGDIQEALRTRQATL
TEIYSQCQRYYQVFQAANDWLEDAQEMLQLAGNGLDVESAEENLKSHMEFFSTEDQFHSNLEELHSLVAT
LDPLIKPTGKEDLEQKVASLELRSQRMSRDSGAQVDLLQRCTAQWHDYQKAREEVIELMNDTEKKLSEFS
LLKTSSSHEAEEKLSEHKALVSVVNSFHEKIVALEEKASQLEKTGNDASKATLSRSMTTVWQRWTRLRAV
AQDQEKILEDAVDEWTGFNNKVKKATEMIDQLQDKLPGSSAEKASKAELLTLLEYHDTFVLELEQQQSAL
GMLRQQTLSMLQDGAAPTPGEEPPLMQEITAMQDRCLNMQEKVKTNGKLVKQELKDREMVETQINSVKCW
VQETKEYLGNPTIEIDAQLEELQILLTEATNHRQNIEKMAEEQKEKYLGLYTILPSELSLQLAEVALDLK
IRDQIQDKIKEVEQSKATSQELSRQIQKLAKDLTTILTKLKAKTDNVVQAKTDQKVLGEELDGCNSKLME
LDAAVQKFLEQNGQLGKPLAKKIGKLTELHQQTIRQAENRLSKLNQAASHLEEYNEMLELILKWIEKAKV
LAHGTIAWNSASQLREQYILHQVTLGKIIFKK

Human SYNE1 Protein sequence - var9 (public gi: 20521662) (SEQ ID NO: 303
WISLMENVIQKDEDNIKNSIGYKAIHEYLQKYKGFKIDINCKQLTVDFVNQSVLQISSQDVESKRSDKTD
FAEQLGAMNKSWQILQGLVTEKIQLLEGLLESWSEYENNVQCLKTWFETQEKRLKQQHRIGDQASVQNAL
KDCQDLEDLIKAKEKEVEKIEQNGLALIQNKKEDVSSIVMSTLRELGQTWANLDHMVGQLKILLKSVLDQ
WSSHKVAFDKINSYLMEARYSLSRFRLLTGSLEAVQVQVDNLQNLQDDLEKQERSLQKFGSITNQLLKEC
HPPVTETLTNTLKEVNMRWNNLLEEIAEQLQSSKALLQLWQRYKDYSKQCASTVQQQEDRTNELLKAATN
KDIADDEVATWIQDCNDLLKGLGTVKDSLFVLHELGEQLKQQVDASAASAIQSDQLSLSQHLCALEQALC
KQQTSLQAGVLDYETFAKSLEALEAWIVEAEEILQGQDPSHSSDLSTIQERMEELKGQMLKFSSMAPDLD
RLNELGYRLPLNDKEIKRMQNLNRHWSLISSQTTERFSKLQSFLLQHQTFLEKCETWMEFLVQTEQKLAV
EISGNYQHLLEQQRAHELFQAEMFSRQQILHSIIIDGQRLLEQGQVDDRDEFNLKLTLLSNQWQGVIRRA
QQRRGIIDSQIRQWQRYREMAEKLRKWLVEVSYLPMSGLGSVPIPLQQARTLFDEVQFKEKVFLRQQGSY
ILTVEAGKQLLLSADSGAEAALQAELAEIQEKWKSASMRLEEQKKKLAFLLKDWEKCEKGIADSLEKLRT
FKKKLSQSLPDHHEELHAEQMRCKELENAVGSWTDDLTQLSLLKDTLSAYISADDISILNERVELLQRQW
EELCHQLSLRRQQIGERLNEWAVFSEKNKELCEWLTQMESKVSQNGDILIEEMIEKLKKDYQEEIAIAQE
NKIQLQQMGERLAKASHESKASEIEYKLGKVNDRWQHLLDLIAARVKKLKETLVAVQQLDKNMSSLRTWL
AHIESELAKPIVYDSCNSEEIQRKLNEQQELQRDIEKHSTGVASVLNLCEVLLHDCDACATDAECDSIQQ
ATRNLDRRWRNICAMSMERRLKIEETWRLWQKFLDDYSRFEDWLKSSERTAAFPSSSGVIYTVAKEELKK
FEAFQRQVHECLTQLELINKQYRRLARENRTDSACSLKQMVHEGNQRWDNLQKRVTSILRRLKHFIGQRE
EFETARDSILVWLTEMDLQLTNIEHFSECDVQAKIKQLKAFQQEISLNHNKIEQIIAQGEQLIEKSEPLD
AAIIEEELDELRRYCQEVFGRVERYHKKLIRLPLPDDEHDLSDRELELEDSAALSDLHWHDRSADSLLSP
QPSSNLSLSLAQPLRSERSGRDTPASVDSIPLEWDHDYDLSRDLESAMSRALPSEDEEGQDDKDFYLRGA
VALSGDHSALESQIRQLGKALDDSRFQIQQTENIIRSKTPTGPELDTSYKGYMKLLGECSSSIDSVKRLE
HKLKEEEESLPGFVNLHSTETQTAGVIDRWELLQAQALSKELRMKQNLQKWQQFNSDLNSIWAWLGDTEE
ELEQLQRLELSTDIQTIELQIKKLKELQKAVDHRKAIILSINLCSPEFTQADSKESRDLQDRLSQMNGRW
DRVCSLLEEWRGLLQDALMQCQGFHEMSHGLLLMLENIDRRKNEIVPIDSNLDAEILQDHHKQLMQIKHE
LLESQLRVASLQDMSCQLLVNAEGTDCLEAKEKVHVIGNRLKLLLKEVSRHIKELEKLLDVSSSQQDLSS
WSSADELDTSGSVSPTSGRSTPNRQKTPRGKCSLSQPGPSVSSPHSRSTKGGSDSSLSEPGPGRSGRGFL
FRVLRAALPLQLLLLLLIGLACLVPMSEEDYSCALSNNFARSFHPMLRYTNGPPPL

Human SYNE1 Protein sequence - var10 (public gi: 28195689) (SEQ ID NO: 304
MTEKLQYLTSVYCTEKMSQQVAELGRETEELRQMIKIRLQNLQDAAKDMKKFEAELKKLQAALEQAQATL
TSPEVGRLSLKEQLSHRQHLLSEMESLKPKVQAVQLCQSALRIPEDVVASLPLCHAALRLQEEASRLQHT
AIQQCNIMQEAVVQYEQYEQEMKHLQQLIEGAHREIEDKPVATSNIQELQAQISRHEELAQKIKGYQEQI
ASLNSKCKMLTMKAKHATMLLTVTEVEGLAEGTEDLDGELLPTPSAHPSVVMMTAGRCHTLLSPVTEESG
EEGTNSEISSPPACRSPSPVANTDASVNQDIAYYQALSAERLQTDAAKIHPSTSASQEFYEPGLEPSATA
KLGDLQRSWETLKNVISEKQRTLYEALERQQKYQDSLQSISTKMEAIELKLSESPEPGRSPESQMAEHQA
LMDEILMLQDEINELQSSLAEELVSESCEADPAEQLALQSTLTVLAERMSTIRMKASGKRQLLEEKLNDQ
LEEQRQEQALQRYRCEADELDSWLLSTKATLDTALSPPKEPMDMEAQLMDCQNMLVEIEQKVVALSELSV
HNENLLLEGKAHTKDEAEQLAGKLRRLKGSLLELQRALHDKQLNMQGTAQEKEESDVDLTATQSPGVQEW
LAQARTTWTQQRQSSLQQQKELEQELAEQKSLLRSVASRGEEILIQHSAAETSGDAGEKPDVLSQELGME
GEKSSAEDQMRMKWESLHQEFSTKQKLLQNVLEQEQEQVLYSRPNRLLSGVPLYKGDVPTQDKSAVTSLL
DGLNQAFEEVSSQSGGAKRQSIHLEQKLYDGVSATSTWLDDVEERLFVATALLPEETETCLFNQEILAKD
IKEMSEEMDKNKNLFSQAFPENGDNRDVIEDTLGCLLGRLSLLDSVVNQRCHQMKERLQQILNFQNDLKV
LFTSLADNKYIILQKLANVFEQPVAEQIEAIQQAEDGLKEFDAGIIELKRRGDELQVEQPSMQELSKLQD
MYDELMMIIGSRRSGLNQNLTLKSQYERALQDLADLLETGQEKMAGDQKIIVSSKEEIQQPLDKHKEYFQ

GLESHMILTVTLFRKIISFAVQKETQFHTELMAQASAVLKRAHKRGVELEYILETWSHLDEDQQELSRQL
EVVESSIPSVGLVEENEDRLIDRITLYQHLKSSLNEYQPKLYQVLDDGKRLLISISCSDLESQLNQLGEC
WLSNTNKMSKELHRLETILKHWTRYQSESADLIHWLQSAKDRLEFWTQQSVTVPQELEMVRDHLNAFLEF
SKEVDAQSSLKSSVLSTGNQLLRLKKVDTATLRSELSRIDSQWTDLLTNIPAVQEKLHQLQMDKLPSRHA
ISEVMSWTSLMENAIQKDEDNIKNSIGYKAIHEYLQKYKGFKIDINCKQLTVDFVNQSVLQISSQDVESK
RSDKTDFAEQLGAMNKSWQILQGLVTEKIQLLEGLLESWSEYENNVQCLKTWFETQEKRLKQQHRIGDQA
SVQNALKDCQDLEDLIKAKDKEVEKIEQNGLALIQTKKEDVSSIVMSTLRELGQTWANLDHMVGQLKILL
KSVLDQWSSHKVAFDKINSYLMEARYSLSRFRLLTGSLEAVQVQVDNLQNLQDDLEKQERSLQKFGSITN
QLLKECHPPVTETLTNTLKEVNMRWNNLLEEIAEQLQSSKALLQLWQRYKDYSKQCASTVQQQEDRTNEL
LKAATNKDIADDEVATWIQDCNDLLKGLGTVKDSLFVLHELGEQLKQQVDASAASAIQSDQLSLSQHLCA
LEQALCKQQTSLQAGVLDYETFAKSLEALEAWIVEAEEILQGQDPSHSSDLSTIQERMEELKGQMLKFSS
MAPDLDRLNELGYRLPLNDKEIKRMQNLNRHWSLISSQTTERFSKLQSFLLQHQTFLEKCETWMEFLVQT
EQKLAVEISGNYQHLLEQQRAHELFQAEMFSRQQILHSIIIDGQRLLEQGQVDDRDEFNLKLTLLSNQWQ
GVIRRAQQRRGIIDSQIRQWQRYREMAEKLRKWLVEVSYLPMSGLGSVPIPLQQARTLFDEVQFKEKVFL
RQQGSYILTVEAGKQLLLSADSGAEAALQAELAEIQEKWKSASMRLEEQKKKLAFLLKDWEKCEKGIADS
LEKLRTFKKKLSQSLPDHHEELHAEQMRCKELENAVGSWTDDLTQLSLLKDTLSAYISADDISILNERVE
LLQRQWEELCHQLSLRRQQIGERLNEWAVFSEKNKELCEWLTQMESKVSQNGDILIEEMIEKLKKDYQEE
IAIAQENKIQLQQMGERLAKASHESKASEIEYKLGKVNDRWQHLLDLIAARVKKLKETLVAVQQLDKNMS
SLRTWLAHIESELAKPIVYDSCNSEEIQRKLNEQQELQRDIEKHSTGVASVLNLCEVLLHDCDACATDAE
CDSIQQATRNLDRRWRNICAMSMERRLKIEETWRLWQKFLDDYSRFEDWLKSSERTAAFPSSSGVIYTVA
KEELKKFEAFQRQVHECLTQLELINKQYRRLARENRTDSACSLKQMVHEGNQRWDNLQKRVTSILRRLKH
FIGQREEFETARDSILVWLTEMDLQLTNIEHFSECDVQAKIKQLKAFQQEISLNHNKIEQIIAQGEQLIE
KSEPLDAAIIEEELDELRRYCQEVFGRVERYHKKLIRLPLPDDEHDLSDRELELEDSAALSDLHWHDRSA
DSLLSPQPSSNLSLSLAQPLRSERSGRDTPASVDSIPLEWDHDYDLSRDLESAMSRALPSEDEEGQDDKD
FYLRGAVALSGDHSALESQIRQLGKALDDSRFQIQQTENIIRSKTPTGPELDTSYKGYMKLLGECSSSID
SVKRLEHKLKEEEESLPGFVNLHSTETQTAGVIDRWELLQAQALSKELRMKQNLQKWQQFNSDLNSIWAW
LGDTEEELEQLQRLELSTDIQTIELQIKKLKELQKAVDHRKAIILSINLCSPEFTQADSKESRDLQDRLS
QMNGRWDRVCSLLEEWRGLLQDALMQCQGFHEMSHGLLLMLENIDRRKNEIVPIDSNLDAEILQDHHKQL
MQIKHELLESQLRVASLQDMSCQLLVNAEGTDCLEAKEKVHVIGNRLKLLLKEVSRHIKELEKLLDVSSS
QQDLSSWSSADELDTSGSVSPTSGRSTPNRQKTPRGKCSLSQPGPSVSSPHSRSTKGGSDSSLSEPGPGR
SGRGFLFRVLRAALPLQLLLLLLIGLACLVPMSEEDYSCALSNNFARSFHPMLRYTNGPPPL

Human SYNE1 Protein sequence - var11 (public gi: 28195677) (SEQ ID NO: 305)
MVVAEDLSALRMAEDGCVDADLPDCNCDVTRARVKKLKETLVAVQQLDKNMSSLRTWLAHIESELAKPIV
YDSCNSEEIQRKLNEQQELQRDIEKHSTGVASVLNLCEVLLHDCDACATDAECDSIQQATRNLDRRWRNI
CAMSMERRLKIEETWRLWQKFLDDYSRFEDWLKSSERTAAFPSSSGVIYTVAKEELKKFEAFQRQVHECL
TQLELINKQYRRLARENRTDSACSLKQMVHEGNQRWDNLQKRVTSILRRLKHFIGQREEFETARDSILVW
LTEMDLQLTNIEHFSECDVQAKIKQLKAFQQEISLNHNKIEQIIAQGEQLIEKSEPLDAAIIEEELDELR
RYCQEVFGRVERYHKKLIRLPLPDDEHDLSDRELELEDSAALSDLHWHDRSADSLLSPQPSSNLSLSLAQ
PLRSERSGRDTPASVDSIPLEWDHDYDLSRDLESAMSRALPSEDEEGQDDKDFYLRGAVALSDVMIPESP
EAYVKLTENAIKNTSGDHSALESQIRQLGKALDDSRFQIQQTENIIRSKTPTGPELDTSYKGYMKLLGEC
SSSIDSVKRLEHKLKEEEESLPGFVNLHSTETQTAGVIDRWELLQAQALSKELRMKQNLQKWQQFNSDLN
SIWAWLGDTEEELEQLQRLELSTDIQTIELQIKKLKELQKAVDHRKAIILSINLCSPEFTQADSKESRDL
QDRLSQMNGRWDRVCSLLEEWRGLLQDALMQCQGFHEMSHGLLLMLENIDRRKNEIVPIDSNLDAEILQD
HHKQLMQIKHELLESQLRVASLQDMSCQLLVNAEGTDCLEAKEKVHVIGNRLKLLLKEVSRHIKELEKLL
DVSSSQQDLSSWSSADELDTSGSVSPTSGRSTPNRQKTPRGKCSLSQPGPSVSSPHSRSTKGGSDSSLSE
PGPGRSGRGFLFRVLRAALPLQLLLLLLIGLACLVPMSEEDYSCALSNNFARSFHPMLRYTNGPPPL

Human SYNE1 Protein sequence - var12 (public gi: 28192628) (SEQ ID NO: 306)
MATSRGASRCPRDIANVMQRLQDEQEIVQKRTFTKWINSHLAKRKPPMVVDDLFEDMKDGVKLLALLEVL
SGQKLPCEQGRRMKRIHAVANIGTALKFLEGRKIKLVNINSTDIADGRPSIVLGLMWTIILYFQIEELTS
NLPQLQSLSSSASSVDSIVSSETPSPPSKRKVTTKIQGNAKKALLKWVQYTAGKQTGIEVKDFGKSWRSG
VAFHSVIHAIRPELVDLETVKGRSNRENLEDAFTIAETELGIPRLLDPEDVDVDKPDEKSIMTYVAQFLK
HYPDIHNASTDGQEDDEILPGFPSFANSVQNFKREDRVIFKEMKVWIEQFERDLTRAQMVESNLQDKYQS
FKHFRVQYEMKRKQIEHLIQPLHRDGKLSLDQALVKQSWDRVTSRLFDWHIQLDKSLPAPLGTIGAWLYR
AEVALREEITVQQVHEETANTIQRKLEQHK

Human SYNE1 Protein sequence - var13 (public gi: 28192522) (SEQ ID NO: 307)
HIQLDKSLPAPLGTIGAWLYRAEVALREEITVQQVHEETANTIQRKLEQHKRKCRTMMDLLQNTDAHKRA
FHEIYRTRSVNGIPVPPDQLEDMAERFHFVSPTSELHLMKMEFLELKYRLLSLLVLAESKLKSWIIKYGR
RESVEQLLQNYVSFIENSKFFEQYEVTYQILKQTAEMYVKADGSVEEAENVMKFMNETTAQWRNLSVEVR
SVRSMLEEVISNWDRYGNTVASLQAWLEDAEKMLNQSENAKKDFFRNLPHWIQQHTAMNDAGNFLIETCD

EMVSRDLKQQLLLLNGRWRELFMEVKQYAQADEMDRMKKEYTDCVVTLSAFATEAHKKLSEPLEVSFMNV
KLLIQDLEDIEQRVPVMDAQYKIITKTAHLITKESPQEEGKEMFATMSKLKEQLTKVKECYSPLLYESQQ
LLIPLEELEKQMTSFYDSLGKINEIITVLEREAQSSALFKQKHQ

Unigene Name: TTC3 Unigene ID: Hs.118174 Clone ID: GD_1105

Human TTC3 mRNA sequence - var1 (public gi: 2687860) (SEQ ID NO: 202)

ATTAAAATAAACATCTTCTGGCCACTTCTGTTTCAACATCAAAACAGTTCCGTAATATCACGATTGCATC
CCTGTGTGGACGCCAACAATTCACGTGCTTCTGAGATAAATTTGAAGAAACTACAACATCTTGAGTTGAT
GGAAGATATTGTGGATTTGGCAAAGAAAGTTGCTAATGATTCATTCCTTATTGGAGGCTTATTGAGAATT
GGTTGTAAAATAGAAAATAAAATCTTGGCAATGGAAGAAGCTCTGAATTGGATAAAATATGCAGGCGATG
TAACAATTCTAACTAAATTAGGATCAATTGACAATTGTTGGCCTATGTTAAGTATTTTCTTTACTGAATA
CAAGTACCACATAACTAAAATTGTAATGGAAGACTGCAATTTGCTTGAAGAACTTAAAACTCAAAGTTGT
ATGGATTGTATAGAGGAAGGAGAACTAATGAAAATGAAAGGAAATGAAGAGTTTTCCAAAGAAAGATTTG
ATATAGCTATTATCTATTACACCAGAGCCATTGAATATAGACCTGAAAACTACCTTCTTTATGGTAACCG
AGCTCTTTGTTTTCTTCGTACTGGACAGTTTAGAAATGCACTCGGTGATGGAAAGAGAGCCACTATTCTG
AAGAACACTTGGCCAAAGGGTCATTATCGTTATTGTGATGCTCTTTCTATGCTGGGGGAATATGACTGGG
CCCTGCAAGCAAACATAAAAGCTCAAAAACTCTGTAAAAATGACCCTGAGGGAATCAAGGATCTAATTCA
GCAGCATGTAAAGTTACAAAAACAAATAGAAGACCTACAAGGTCGAACAGCAAATAAGGATCCAATTAAA
GCCTTTTATGAAAACAGGGCCTACACACCTAGGAGTTTATCAGCACCTATATTTACTACTTCACTTAACT
TTGTGGAGAAGGAAAGAGATTTCAGAAAAATTAATCACGAAATGGCCAACGGTGGTAATCAGAATCTAAA
GGTGGCGGATGAGGCGTTGAAGGTAGATGATTGTGACTGTCATCCTGAATTTTCACCACCATCAAGTCAG
CCTCCAAAACATAAAGGAAAACAAAAATCTCGAAACAATGAATCAGAAAAGTTCAGTTCTAGTTCACCAT
TGACTTTACCAGCAGATTTGAAGAACATCTTGGAGAAACAGTTTTCTAAATCTTCAAGAGCTGCACACCA
GGATTTTGCTAATATAATGAAAATGCTGAGAAGCTTAATTCCAGATGGCTATATGGCCTTATTGGAGCAG
CGTTGCCGCAGCGCTGCACAGGCCTTTACAGAGTTGCTGAACGGTTTAGATCCTCAAAAAATAAAGCAAT
TGAACCTGGCCATGATTAACTATGTTTTGGTCGTCTATGGACTTGCCATTTCTCTCCTTGGAATAGGACA
GCCTGAGGAATTATCTGAAGCCGAAAACCAGTTTAAGAGGATTATTGAACACTACCCCAGTGAGGGCCTT
GATTGCTTGGCCTACTGTGGAATTGGAAAAGTATATTTGAAAAAAAACAGATTTCTAGAAGCTCTCAATC
ACTTTGAGAAAGCAAGAACCTTGATTTATCGTCTTCCTGGAGTGTTAACTTGGCCCACGAGTAATGTGAT
TATTGAAGAGTCTCAGCCACAAAAAATAAAGATGCTGTTAGAGAAATTTGTTGAAGAATGCAAGTTCCCT
CCAGTGCCAGATGCCATTTGTTGCTATCAGAAGTGCCATGGATATTCTAAGATCCAGATATACATAACTG
ATCCAGACTTTAAGGGTTTTATACGCATCAGCTGTTGCCAGTACTGTAAAATAGAATTTCACATGAATTG
CTGGAAGAAGTTAAAAACTACAACCTTTAATGATAAAATTGACAAGGATTTTCTACAAGGAATATGTCTT
ACCCCTGACTGTGAAGGTGTCATTTCTAAGATTATCATCTTCAGCAGTGGTGGTGAAGTTAAATGTGAAT
TTGAACACAAGGTCATAAAAGAAAAGGTTCCTCCAAGACCTATTCTGAAACAGAAATGTTCTAGCCTAGA
GAAACTAAGACTGAAAGAAGACAAAAAATTGAAGAGAAAGATCCAAAAAAAAGAAGCAAAAAAGTTAGCA
CAAGAAAGAATGGAGGAGGACTTAAGAGAAAGTAATCCACCCAAAAATGAAGAGCAGAAAGAAACTGTAG
ACAATGTTCAGCGTTGTCAGTTCCTTGATGACAGAATTCTACAGTGTATAAAGCAGTATGCTGACAAGAT
TAAATCCGGCATACAGAATACAGCCACGCTTCTCAAAGAATTGCTTTCTTGGAAAGTTTTGAGCACAGAA
GACTATACAACCTGTTTTTCTAGCAGAAATTTTCTAAATGAAGCAGTGGACTATGTTATTCGCCACTTGA
TTCAAGAAAATAACAGAGTAAAGACAAGAATATTTCTGCATGTTTTGAGTGAGCTTAAAGAAGTGGAGCC
CAAATTAGCCGCCTGGATCCAAAAACTTAATAGCTTTGGCTTAGATGCCACAGGAACTTTCTTTTCTCGT
TATGGAGCATCTCTTAAACTGCTTGATTTTAGTATCATGACTTTCCTCTGGAATGAGAAATATGGTCACA
AACTAGACTCTATAGAAGGAAAGCAACTTGATTATTTCTCTGAGCCAGCATCATTGAAGGAAGCCCGTTG
TTTAATATGGCTGCTAGAAGAACACAGAGACAAGTTCCCAGCATTGCATAGTGCTTTAGATGAATTCTTT
GATATAATGGACAGCCGCTGTACTGTGTTAAGGAAACAAGATAGTGGTGAAGCACCGTTTAGTTCAACCA
AGGTGAAAAACAAAAGCAAGAAAAAGAAGCCAAAGGATTCAAAGCCTATGTTAGTTGGGTCTGGAACAAC
TTCAGTAACTTCAAATAATGAGATCATCACTTCAAGTGAAGACCATAGCAATCGAAATTCAGATTCTGCA
GGCCCATTTGCAGTGCCTGACCATCTTCGGCAAGATGTAGAAGAATTCGAAGCTCTCTATGACCAACACA
GTAACGAATATGTTGTCCGCAATAAGAAGCTATGGGACATGAACCCAAAACAAAAATGTTCAACTCTATA
TGATTACTTCTCTCAGTTTTTGGAGGAACATGGTCCCTTGGACATGAGTAACAAGATGTTCTCTGCAGAA
TATGAGTTTTTCCCAGAAGAAACTCGACAGATACTAGAAAAAGCAGGAGGTTTAAAACCTTTTCTCTTGG
GATGCCCTCGTTTTGTTGTGATTGACAACTGTATTGCACTGAAGAAGGTTGCATCACGGCTCAAGAAAAA
AAGGAAGAAGAAAAACATTAAAACAAAAGTAGAAGAAATTTCAAAAGCAGGGGAGTATGTACGAGTTAAA
CTACAACTGAATCCAGCTGCTAGGGAATTTAAACCAGATGTAAAGTCTAAACCAGTGTCAGATTCATCTT
CAGCACCAGCTTTTGAAAATGTGAAACCCAAACCTGTGTCTGCAAATTCTCCCAAGCCAGCTTGTGAAGA
TGTGAAGGCCAAACCAGTATCCGACAATTCTTCTAGACAAGTTTCTGAGGATGGGCAACCCAAAGGGGTC
TCTTCTAATTCTCCTAAACCAGGCTCTGAGGATGCAAATTACAAGCGAGTCTCCTGTAATTCCCCCAAAC

CGGTTCTTGAGGATGTGAAACCAACTTATTGGGCTCAATCCCATTTGGTCACAGGATACTGTACGTATCT
TCCTTTCCAGAGATTTGATATCACCCAGACACCGCCAGCATACATAAACGTGTTACCAGGTTTGCCCCAG
TACACCAGCATATATACACCCTTGGCCAGCCTTTCTCCTGAATATCAGCTACCAAGATCAGTACCAGTGG
TGCCGTCTTTTGTAGCCAATGACAGAGCAGATAAAAATGCTGCTGCCTATTTTGAGGGTCATCATTTGAA
TGCTGAGAATGTTGCTGGTCACCAGATTGCCTCTGAAACACAGATCCTTGAGGGCTCTTTGGGAATATCT
GTAAAGTCACACTGCAGCACAGGTGATGCTCATACAGTCCTGAGTGAGTCTAACAGAAATGATGAGCACT
GTGGAAATTCTAACAACAAATGTGAAGTAATTCCAGAAAGCACCAGTGCAGTAACAAACATTCCACACGT
GCAGATGGTTGCCATACAGGTATCTTGGAACATAATACACCAAGAAGTCAATACTGAGCCATATAATCCT
TTTGAGGAACGACAAGGGGAAATTTCACGGATTGAAAAGGAGCACCAAGTATTACAAGACCAACTTCAAG
AAGTGTATGAAAATTATGAGCAGATAAAACTTAAGGGCTTAGAAGAGACCAGGGACCTGGAAGAGAAGTT
GAAAAGGCACTTAGAAGAAAACAAGATCTCAAAGACGGAATTAGATTGGTTCCTTCAAGATTTGGAAAGA
GAAATTAAAAAATGGCAACAGGAAAAAAAAGAAATCCAAGAAAGACTAAAATCACTGAAGAAGAAAATTA
AAAAGGTTTCAAATGCCAGTGAAATGTATACCCAGAAAAATGATGGAAAGGAAAAGGAACATGAATTACA
TCTGGATCAGTCCCTTGAAATCAGCAACACACTTACAAATGAGAAAATGAAAATAGAAGAGTATATAAAG
AAAGGGAAAGAGGATTATGAAGAGAGTCATCAGAGAGCTGTGGCTGCAGAGGTATCCGTACTTGAAAACT
GGAAGGAGAGTGAAGTGTATAAGCTACAGATCATGGAGTCACAAGCAGAAGCCTTTCTGAAGAAGCTGGG
GCTGATTAGCCGTGATCCTGCAGCATATCCTGACATGGAGTCTGATATACGTTCATGGGAATTGTTTCTT
TCTAATGTTACAAAAGTAATTGAGAAAGCAAAGTCTCAGTTTGAAGAACAAATTAAGGCAATTAAAAATG
GTTCTCGGCTCAGTGAACTTTCTAAAGTGCAGATTTCTGAGCTTTCATTTCCTGCCTGTAACACGGTTCA
TCCCGAGTTACTCCCTGAGTCTTCAGGCGACGATGGCCAAGGGCTTGTGACTTCTGCAAGCGACGTGACT
GGAAACCACGCAGCACTTCACAGGGATCCTAGTGTGTTCTCTGCTGGTGATTCCCCAGGGGAGGCTCCTT
CTGCGCTGTTGCCAGGGCCACCCCCTGGTCAGCCTGAAGCCACTCAGCTGACAGGGCCAAAACGGGCTGG
CCAGGCAGCTCTGTCAGAACGAAGCCCTGTGACTGATCGGAAGCAGCCTGTTCCTCCAGGACGTGCTGCG
CGTTCAAGCCAGTCTCCAAAAAAGCCGTTCAATAGTATTATTGAGCACCTGTCAGTGGTATTCCCATGTT
ACAACAGCACTGAGCTTGCTGGTTTTATTAAAAAAGTGCGAAGCAAAAACAAGAACTCACTCTCAGGATT
GAGTATTGATGAAATTGTCCAAAGAGTGACAGAACACATTCTAGATGAACAGAAAAAGAAAAAGCCAAAC
CCAGGAAAGGACAAGAGGACTTATGAGCCCAGCTCTGCCACCCCCGTGACCAGGTCCTCCCAGGGCTCAC
CCTCGGTGGTTGTTGCACCATCACCCAAAACCAAGGGGCAGAAAGCAGAAGATGTCCCTGTGAGGATTGC
ACTGGGTGCAAGTTCCTGTGAAATATGCCACGAGGTGTTCAAATCAAAAAACGTGCGTGTGCTCAAATGT
GGGCACAAGTATCACAAAGGGTGCTTTAAGCAGTGGCTTAAAGGGCAGAGCGCTTGCCCGGCCTGCCAGG
GTCGTGATCTCCTGACAGAAGAGTCACCTTCTGGAAGAGGCTGGCCCAGTCAGAATCAGGAGCTGCCTTC
CTGCTCTTCTAGGTAGTCACACTTCACTAAAGTGTCATCCACCAGTGTGTTGAATCCGAAGAATGACAAT
TTTCTACCACTGGTGTAAAAAACAAACATTTGAAGACCCTTGTGCATTGTGTGTCACAAAGCTAAATACA
TGGAAATCGTTAATATCGCTGATATTAAGTAATTTCCCCACTCTGAGTGAATACTTTGATGATTGCCAAC
AGTGGCTAATAAAATGACGGCTACCACACTCATGGGTCACTGGGGCTGCGCAGGGCTCTTTGAGGTGGGT
GGCTTCTTTTGGAAAGTACTATGAACGTCTCGAAGCAGTATTCTAGTGATAAGAATTCTTAACATAGCCA
AGCGCCCCACGTTTGTTCCCCACGTTTGTTCCCCTTTTCTGTTTGAAAAACCTGTTCTGGTAGCTCCACA
AGAGAGATGATACTGACTTTTTAAATTTTTTACAAGAGTCTGTATTCCTGATATGCCTATATTTTTCCTC
AAAGATTCTGCATTTTAAGGATGGGCATAAGCAAACTATATTTTAATAATTTATAGTTAATGTTAAAATA
TTGGCTGATTTAGACCAAAAGATTCAAATCTCCTCTTTGTGAAATCCCATCTGCATTTGATTTTTTATTA
TTTTATGTTCCCCCGTTAGATTGTTTTAAGTGTTTGCTTTTCATCTTTTATAGATGTAATCTGATTTTCA
AAAATCATTAACACTTTTTAATTAGTATCGACTAAGACTTTTTCCCCCTGGAATCGAGGCTGTGTGTCCG
TCATCCCAGCCCCCGGTTGGAGCCTGCTCTTTGAACTCCGCTGCCTTCCTTAGCAGCTTCTGTCCTCTTC
TGTGAGTCAGTCAGCGAGTGCTTGGGATCCGCATCCAGCCGTGCTGAGCACACAACAGGCTGTGTGTGGA
AATGGCCACCACCATTCTCCTTCCCCACCCCACCACAAAAAGAGAAGCTGTGTCTTTAGACAACCCTGAG
GTATCTGTGTTACAATCGTTCTGTGTTTGATATTTGTGTAAAGTATGCATGCAGTCTTGTACTGTGACCT
AAGAACAAAACTGTAACTGCATTAGAAACCATGAAAAAATTAGATATTGTTTTGTGACTTTTAGACAGTG
GTAAATATAGAACCATGAATTCTGGTCACATTCCATTTCTCTCCAACATGAAGGATCAAAAAATGTTTTT
CAATGTGTTCTTTGTTCCACTGGAAACTTAGAGTCATGAGTTTATGAGCTGATTTGGTCACCTTCCTCTG
CCTTTGTTCACTGTGAGTTCTGATGTCTTAGTGACTTAGTTCTTAGAAGCTCACGCCTTAGTTTGAAACA
GATTCTCCACGGTGGTCCCCAAAACACTGTCTGCATATCCATAAGAATTGAGCGCTATGGGTGTTAACGT
GCATGAGGATCAGTTTGCAGCAGCAAGTACAAAAGGAGAAGAGGAACATCCGTTGAATGAGTGTGTTTTG
TACATAACTTCAGATACTTGTGAACATGCCTTATATTTGTCCAACAACTGTCAGAATAAAGAACATTCT

Human TTC3 mRNA sequence - var2 (public gi: 1632765) (SEQ ID NO: 203)
TACATTTGGAAAGTCTTACTGACATGCAGAAATAGTACAGAAAAACATACAAATAGGAATGTTATTGGCT
GGGCATGGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATCACAAGGTCAGGA
GATCGAGACCATCGTGGCTAACATGGTGAAACCCTGTGCCTACTAAAAATTCAAAAAATTAGCCAGGTGT
ACTGGCATGTGCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCACTTGAACCCGGGAGCAAA
GGTTGCAGTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCTGTCTCAAAAA
AAACAAAAGAATGCTATGCATAGGTACAATGTCGAAATGTGCAAGAAATACTTCAGAAATATTAAAAGTA
GTTATTCCTGGGTAGTTGTGATTGTGACTCTGGGTGACTTTTTCCCTTTGTTTTATTTTTCTGTATTTTC
CAAATTTCCTATAATGGACATATATTATATGGATTTTTTTAATTAAAATTATCTTTTGACTAGATAATA

TACATGGAAAATTCACAAAGTACACATATGCAAGATATAATACTTTTACTATAAAAGAGTGAAAATTTTT
AAATACTTGTTTTTCCTTTTAGAAATGCACTCGGTGATGGAAAGAGAGCCACTATTCTGAAGAACACTTG
GCCAAAGGGTCATTATCGTTATTGTGATGCTCTTTCTATGCTGGGGAATATGACTGGGCCCTGCAAGCA
AACATAAAAGCTCAAAAACTCTGTAAAAATGACCCTGAGGGAATCAAGGATCTAATTCAGCAGCATGTAA
AGTTACAAAAACAAATAGAAGACCTACAAGGTCGAACAGCAAATAAGGATCCAATTAAAGCCTTTTATGA
AAACAGGGCCTACACACCTAGGAGTTTATCAGCACCTATATTTACTACTTCACTTAACTTTGTGGAGAAG
GAAAGAGATTTCAGAAAAATTAATCACGAAATGGCCAACGGTGGTAATCAGAATCTAAAGGTGGCGGATG
AGGCGTTGAAGGTAGATGATTGTGACTGTCATCCTGAATTTTCACCACCATCAAGTCAGCCTCCAAAACA
TAAAGGAAAACAAAAATCTCGAAACAATGAATCAGAAAAGTTCAGTTCTAGTTCACCATTGACTTTACCA
GCAGATTTGAAGAACATCTTGGAGAAACAGTTTTCTAAATCTTCCAGAGCTGCACACCAGGATTTTGCTA
ATATAATGAAAATGCTGAGAAGCTTAATTCAAGATGGCTATATGGCCTTATTGGAGCAGCGTTGCCGCAG
CGCTGCACAGGCCTTTACAGAGTTGCTGAACGGTTTAGATCCTCAAAAAATAAAGCAATTGAACCTGGCC
ATGATTAACTATGTTTTGGTCGTCTATGGACTTGCCATTTCTCTCCTTGGAATAGGACAGCCTGAGGAAT
TATCTGAAGCCGAAAACCAGTTTAAGAGGATTATTGAACACTACCCCAGTGAGGGCCTTGATTGCTTGGC
CTACTGTGGAATTGGAAAAGTGTATTTGAAAAAAAACAGATTTCTAGAAGCTCTCAATCACTTTGAGAAA
GCAAGAACCTTGATTTATCGTCTTCCTGGAGTGTTAACTTGGCCCACGAGTAATGTGATTATTGAAGAGT
CTCAGCCACAAAAAATAAAGATGCTGTTAGAGAAATTTGTTGAAGAATGCAAGTTCCCTCCAGTGCCAGA
TGCCATTTGTTGCTATCAGAAGTGCCATGGATATTCTAAGATCCAGATATACATAACTGATCCAGACTTT
AAGGGTTTTATACGCATCAGCTGTTGCCAGTACTGTAAAATAGAATTTCACATGAATTGCTGGAAGAAGT
TAAAAACTACAACCTTTAATGATAAAATTGACAAGGATTTTCTACAAGGAATATGTCTTACCCCTGACTG
TGAAGGTGTCATTTCTAAGATTATCATCTTCAGCAGTGGTGGTGAAGTTAAATGTGAATTTGAACACAAG
GTCATAAAAGAAAAGGTTCCTCCAAGACCTATTCTGAAACAGAAATGTTCTAGCCTAGAGAAACTAAGAC
TGAAAGAAGACAAAAAATTGAAGAGAAAGATCCAAAAAAAAGAAGCAAAAAAGTTAGCACAAGAAAGAAT
GGAGGAGGACTTAAGAGAAAGTAATCCACCCAAAAATGAAGAGCAGAAAGAAACTGTAGACAATGTTCAG
CGTTGTCAGTTCCTTGATGACAGAATTCTACAGTGTATAAAGCAGTATGCTGACAAGATTAAATCCGGCA
TACAGAATACAGCCATGCTTCTCAAAGAATTGCTTTCTTGGAAAGTTTTGAGCACAGAAGACTATACAAC
CTGTTTTTCTAGCAGAAATTTTCTAAATGAAGCAGTGGACTATGTTATTCGCCACTTGATTCAAGAAAAT
AACAGAGTAAAGACAAGAATATTTCTGCATGTTTTGAGTGAGCTTAAAGAAGTGGAGCCCAAATTAGCCG
CCTGGATCCAAAAACTTAATAGCTTTGGCTTAGATGCCACAGGAACTTTCTTTTCTCGTTATGGAGCATC
TCTTAAACTGCTTGATTTTAGTATCATGACTTTCCTCTGGAATGAGAAATATGGTCACAAACTAGACTCT
ATAGAAGGAAAGCAACTTGATTATTTCTCTGAGCCAGCATCATTGAAGGAAGCCCGTTGTTTAATATGGC
TGCTAGAAGAACACAGAGACAAGTTCCCAGCATTGCATAGTGCTTTAGATGAATTCTTTGATATAATGGA
CAGCCGCTGTACTGTGTTAAGGAAACAAGATAGTGGTGAAGCACCGTTTAGTTCAACCAAGGTGAAAAAC
AAAAGCAAGAAAAAGAAGCCAAAGGATTCAAAGCCTATGTTAGTTGGGTCTGGAACAACTTCAGTAACTT
CAAATAATGAGATCATCACTTCAAGTGAAGACCATAGCAATCGAAATTCAGATTCTGCAGGCCCATTTGC
AGTGCCTGACCATCTTCGGCAAGATGTAGAAGAATTCGAAGCTCTCTATGACCAACACAGTAACGAATAT
GTTGTCCGCAATAAGAAGCTATGGGACATGAACCCAAAACAAAAATGTTCAACTCTATATGATTACTTCT
CTCAGTTTTTGGAGGAACATGGTCCCTTGGACATGAGTAACAAGATGTTCTCTGCAGAATATGAGTTTTT
CCCAGAAGAAACTCGACAGATACTAGAAAAAGCAGGAGGTTTAAAACCTTTTCTCTTGGGATGCCCTCGT
TTTGTTGTGATTGACAACTGTATTGCACTGAAGAAGGTTGCATCACGGCTCAAGAAAAAAAAGGAAGAAGA
AAAACATTAAAACAAAAGTAGAAGAAATTTCAAAAGCAGGGGAGTATGTACGAGTTAAACTACAACTGAA
TCCAGCTGCTAGGGAATTTAAACCAGATGTAAAGTCTAAACCAGTGTCAGATTCATCTTCAGCACCAGCT
TTTGAAAATGTGAAACCCAAACCTGTGTCTGCAAATTCTCCCAAGCCAGCTTGTGAAGATGTGAAGGCCA
AACCAGTATCCGACAATTCTTCTAGACAAGTTTCTGAGGATGGCAACCCAAAGGGGTCTCTTCTAATTC
TCCTAAACCAGGCTCTGAGGATGCAAATTACAAGCGAGTCTCCTGTAATTCCCCCAAACCGGTTCTTGAG
GATGTGAAACCAACTTATTGGGCTCAATCCCATTTGGTCACAGGATACTGTACGTATCTTCCTTTCCAGA
GATTTGATATCACCCAGACACCGCCAGCATACATAAACGTGTTACCAGGTTTGCCCCAGTACACCAGCAT
ATATACACCCTTGGCCAGCCTTTCTCCTGAATATCAGCTACCAAGATCAGTACCAGTGGTGCCGTCTTTT
GTAGCCAATGACAGAGCAGATAAAAATGCTGCTGCCTATTTTGAGGGTCATCATTTGAATGCTGAGAATG
TTGCTGGTCACCAGATTGCCTCTGAAACACAGATCCTTGAGGGCTCTTTGGGAATATCTGTAAAGTCACA
CTGCAGCACAGGTGATGCTCATACAGTCCTGAGTGAGTCTAACAGAAATGATGAGCACTGTGGAAATTCT
AACAACAAATGTGAAGTAATTCCAGAAAGCACCAGTGCAGTAACAAACATTCCACACGTGCAGATGGTTG
CCATACAGGTATCTTGGAACATAATACACCAAGAAGTCAATACTGAGCCATATAATCCTTTTGAGGAACG
ACAAGGGGAAATTTCACGGATTGAAAAGGAGCACCAAGTATTACAAGACCAACTTCAAGAAGTGTATGAA
AATTATGAGCAGATAAAACTTAAGGGCTTAGAAGAGACCAGGGACCTGGAAGAGAAGTTGAAAAGGCACT
TAGAAGAAAACAAGATCTCAAAGACGGAATTAGATTGGTTCCTTCAAGATTTGGAAAGAGAAATTAAAAA
ATGGCAACAGGAAAAAAAAGAAATCCAAGAAAGACTAAAATCACTGAAGAAGAAAATTAAAAAGGTTTCA
AATGCCAGTGAAATGTATACCCAGAAAAATGATGGAAAGGAAAAGGAACATGAATTACATCTGGATCAGT
CCCTTGAAATCAGCAACACACTTACAAATGAGAAAATGAAAATAGAAGAGTATATAAAGAAAGGGAAAGA
GGATTATGAAGAGAGTCATCAGAGAGCTGTGGCTGCAGAGGTATCCGTACTTGAAAACTGGAAGGAGAGT
GAAGTGTATAAGCTACAGATCATGGAGTCACAAGCAGAAGCCTTTCTGAAGAAGCTGGGGCTGATTAGCC
GTGATCCTGCAGCATATCCTGACATGGAGTCTGATATACGTTCATGGGAATTGTTTCTTTCTAATGTTAC
AAAAGAAATTGAGAAAGCAAAGTCTCAGTTTGAAGAACAAATTAAGGCAATTAAAAATGGTTCTCGGCTC

AGTGAACTTTCTAAAGTGCAGATTTCTGAGCTTTCATTTCCTGCCTGTAACACGGTTCATCCCGAGTTAC
TCCCTGAGTCTTCAGGCCACGATGGCCAAGGGCTTGTGACTTCTGCAAGCGACGTGACTGGAAACCACGC
AGCACTTCACAGGGATCCTAGTGTGTTCTCTGCTGGTGATTCCCCAGGGGAGGCTCCTTCTGCGCTGTTG
CCAGGGCCACCCCCTGGTCAGCCTGAAGCCACTCAGCTGACAGGGCCAAAACGGGCTGGCCAGGCAGCTC
TGTCAGAACGAAGCCCTGTGGCTGATCGGAAGCAGCCTGTTCCTCCAGGACGTGCTGCGCGTTCAAGCCA
GTCTCCAAAAAAGCCGTTCAATAGTATTATTGAGCACCTGTCAGTGGTATTCCCATGTTACAACAGCACT
GAGCTTGCTGGTTTTATTAAAAAAGTGCGAAGCAAAAACAAGAACTCACTCTCAGGATTGAGTATTGATG
AAATTGTCCAAAGAGTGACAGAACACATTCTAGATGAACAGAAAAAGAAAAAGCCAAACCCAGGAAAGGA
CAAGAGGACTTATGAGCCCAGCTCTGCCACCCCCGTGACCAGGTCCTCCCAGGGCTCACCCTCGGTGGTT
GTTGCACCATCACCCAAAACCAAGGGGCAGAAAGCAGAAGATGTCCCTGTGAGGATTGCACTGGGTGCAA
GTTCCTGTGAAATATGCCACGAGGTGTTCAAATCAAAAAACGTGCGTGTGCTCAAATGTGGGCACAAGTA
TCACAAAGGGTGCTTTAAGCAGTGGCTTAAAGGGCAGAGCGCTTGCCCGGCCTGCCAGGGTCGTGATCTC
CTGACAGAAGAGTCACCTTCTGGAAGAGGCTGGCCCAGTCAGAATCAGGAGCTGCCTTCCTGCTCTTCTA
GGTAGTCACACTTCACTAAAGTGTCATCCACCAGTGTGTTGAATCCGAAGAATGACAATTTTCTACCACT
GGTGTAAAAAACAAACATTTGAAGACCCTTGTGCATTGTGTGTCACAAAGCTAAATACATGGAAATCGTT
AATATCGCTGATATTAAGTAATTTCCCCACTCTGAGTGAATACTTTGATGATTGCCAACAGTGGCTAATA
AAATGACGGCTACCACACTCATGGGTCACTGGGGCTGCGCAGGGCTCTTTGAGGTGGGTGGCTTCTTTTG
GAAAGTACTATGAACGTCTCGAAGCAGTATTCTAGTGATAAGAATTCTTAACATAGCCAAGCGCCCCACG
TTTGTTCCCCACGTTTGTTCCCCTTTTCTGTTTGAAAAACCTGTTCTGGTAGCTCCACAAGAGAGATGAT
ACTGACTTTTTAAATTTTTTACAAGAGTCTGTATTCCTGATATGCCTATATTTTTCCTCAAAGATTCTGC
ATTTTAAGGATGGGCATAAGCAAACTATATTTTAATAATTTATAGTTAATGTTAAAATATTGGCTGATTT
AGACCAAAAGATTCAAATCTCCTCTTTGTGAAATCCCATCTGCATTTGATTTTTTATTATTTTATGTTCC
CCCGTTAGATTGTTTTAAGTGTTTGCTTTTCATCTTTTATAGATGTAATCTGATTTTCAAAAATCATTAA
CACTTTTTAATTAGTATCGACTAAGACTTTTTCCCCCTGGAATCGAGGCTGTGTGTCCGTCATCCCAGCC
CCCGGTTGGAGCCTGCTCTTTGAACTCCGCTGCCTTCCTTAGCAGCTTCTGTCCTCTTCTGTGAGTCAGT
CAGCGAGTGCTTGGGATCCGCATCCAGCCGTGCTGAGCACACAACAGGCTGTGTGTGGAAATGGCCACCA
CCATTCTCCTTCCCCACCCCACCACAAAAAGAGAAGCTGTGTCTTTAGACAACCCTGAGGTATCTGTGTT
ACAATCGTTCTGTGTTTGATATTTGTGTAAAGTATGCATGCAGTCTTGTACTGTGACCTAAGAACAAAAC
TGTAACTGCATTAGAAACCATGAAAAAATTAGATATTGTTTTGTGACTTTTAGACAGTGGTAAATATAGA
ACCATGAATTCTGGTCACATTCCATTTCTCTCCAACATGAAGGATCAAAAAATGTTTTTCAATGTGTTCT
TTGTTCCACTGGAAACTTAGAGTCATGAGTTTATGAGCTGATTTGGTCACCTTCCTCTGCCTTTGTTCAC
TGTGAGTTCTGATGTCTTAGTGACTTAGTTCTTAGAAGCTCACGCCTTAGTTTGAAACAGATTCTCCACG
GTGGTCCCCAAAACACTGTCTGCATATCCATAAGAATTGAGCGCTATGGGTGTTAACGTGCATGAGGATC
AGTTTGCAGCAGCAAGTACAAAAGGAGAAGAGGAACATCCGTTGAATGAGTGTGTTTTGTACATAACTTC
AGATACTTGTGAACATGCCTTATATTTGTCCAACAACTGTCAGAATAAAGAACATTCTAAAATGAG

Human TTC3 mRNA sequence - var3 (public gi: 1632763) (SEQ ID NO: 204)
CTGAACTAGTTGCCAGTGATCTTGAAACGTGACAGTAACCAAGAGATAAATAGGTGACAATGACAGGAAA
ATTAGATGTAGTAAAAGAGAGTGTTTGAGAGCAGAAGCTATGGCAACTAAAGACTGGATTTGAATCCTTC
CTAGCTTGGTGACATGAGCAAATTACTTGATTTAAGTGAGCATTTTCCCATCTGTACAGTGGAGATAACG
ATAATTGTGCCTGCTAAGAAGAATTGCTGTGAAGATTAGTGAAATAATGCATGTAAAACATTTGGTACAG
TATGTGACACATAGTACAAATAGTTTGCTAGGAAGATTGTTATTATTCTTCACTTGTGATATTGTGAAGT
TTTCATACAGCAAATTGGACATCATGAGATGGATTGATTAAATAAATAGATTTGAACTTCAAGGACTGGT
AGTGTTCTTGCTTTGGAAAGAAGAAACTTGGTTTATCCTAATAATAGTAGGATAATAATGGTGAAGTGAT
AGGTACAAGTAATAGTGTTTATGATGCGCTGGTGATGATAGGAAAAGAAAGCCATTATATGGGCAAGAGC
TAGAAGTAATAAAATGGTGCATTTTTCAGTGATGTTTGGCCTATGTAGCTATTCTCTGATAACTATAAAA
ATCCTTATTATTGAAGATTCTTCAGGAAAAAAAAAACCCTTAGTCTGAAACTTTAGCACCAATCCCCCTTG
CCCCCCATTGAAATACGTATTTTTAAAACATGGCTTTTGATAATGTGAGGGTTTTTTCCTTTTTTGCGATT
TAGCAGTGCTGATTGTGTATTGCAGTAGTTGTGAGAGCATTAGAAGCAGCAGTCGATAGGAGGATGGAAG
GTCTGGATGCCGCCTTGGGGAGTTAGGAGATTGGCAGACTTACCCTGTACCACTCTAGCCCTACTCCTTT
GCCCAAGACAGAAACACACTGAGATGGATAGGAGAATATGAGCAGTTGATAGGAAAGTTCTCAGTGGAGT
CAGGATTTAGGTTAGGCCAGGAGATTGAGAATATAACAGTTTGTGTATGATGAAATGGCATATTTCACAG
AATGCAGTAAAAGCAGGTAGGGTACAAGTGCAGCAACAGGAAGATGTCTTTTCTTCATTCAGCAAACACT
TATTTGAGAGCTTACCATGTGCTAGGCACATACAAAGATAAATAAGATGCCCTTGATGATCCTCTATTTA
AAGGAGACATGTAAACAGGTTAACTTAGAGTAGAGATGGTGAATATGTGAACCTGAGGAAAGGAAGAAAT
AGATTAAATTATCTGGAGAGAGAGGAAAAGTCAGCAGAATGGGGACGAGAATCTTTCGGAGCTCAGTGTT
CTGATAGGAGTTATTTCCTTGGGCATAGGTTCCAAGTATTTTTCTAATATACCATAGAAGCCAGGAAAAC
TTTCTTCTGTTATCTCAAATGATTTAATTACTGACTTGAGTTTGTGTTGTCTCCTTAGACTTGTGCACCA
TGGACAATTTTGCTGAGGGAGATTTCACTGTGGCGGATTATGCCTTGTTAGAAGATTGCCCTCACGTGGA
TGATTGTGTCTTTGCTGCTGAATTTATGAGCAATGATTATGTTCGTGTGACTCAGCTTTACTGTGATGGG
GTGGGTGTGCAATATAAAGATTATATCCAAAGTGAGAGGAATTTGGAATTTGACATCTGCAGTATATGGT
GTAGTAAACCAATTTCTGTCCTGCAAGATTATTGCGATGCCATTAAAATAAACATCTTCTGGCCACTTCT
GTTTCAACATCAAAACAGTTCCGTAATATCACGATTGCATCCCTGTGTGGACGCCAACAATTCACGTGCT

TCTGAGATAAATTTGAAGAAACTACAACATCTTGAGTTGATGGAAGATATTGTGGATTTGGCAAAGAAAG
TTGCTAATGATTCATTCCTTATTGGAGGCTTATTGAGAATTGGTTGTAAAATAGAAAATAAAATCTTGGC
AATGGAAGAAGCTCTGAATTGGATAAAATATGCAGGCGATGTAACAATTCTAACTAAATTAGGATCAATT
GACAATTGTTGGCCTATGTTAAGTATTTTCTTTACTGAATGAAGGAGAACTAATGAAAATGAAAGGAAAT
GAAGAGTTTTCCAAAGAAAGATTTGATATAGCTATTATCTATTACACCAGAGCCATTGAATATAGACCTG
AAAACTACCTTCTTTATGGTAACCGAGCTCTTTGTTTTCTTCGTACTGGACAGTTTAGAAATGCACTCGG
TGATGGAAAGAGAGCCACTATTCTGAAGAACACTTGGCCAAAGGGTCATTATCGTTATTGTGATGCTCTT
TCTATGCTGGGGGAATATGACTGGGCCCTGCAAGCAAACATAAAAGCTCAAAAACTCTGTAAAAATGACC
CTGAGGGAATCAAGGATCTAATTCAGCAGCATGTAAAGTTACAAAAACAAATAGAAGACCTACAAGGTCG
AACAGCAAATAAGGATCCAATTAAAGCCTTTTATGAAAACAGGGCCTACACACCTAGGAGTTTATCAGCA
CCTATATTTACTACTTCACTTAACTTTGTGGAGAAGGAAAGAGATTTCAGAAAAATTAATCACGAAATGG
CCAACGGTGGTAATCAGAATCTAAAGGTGGCGGATGAGGCGTTGAAGGTAGATGATTGTGACTGTCATCC
TGAATTTTCACCACCATCAAGTCAGCCTCCAAAACATAAAGGAAAACAAAAATCTCGAAACAATGAATCA
GAAAAGTTCAGTTCTAGTTCACCATTGACTTTACCAGCAGATTTGAAGAACATCTTGGAGAAACAGTTTT
CTAAATCTTCCAGAGCTGCACACCAGGATTTTGCTAATATAATGAAAATGCTGAGAAGCTTAATTCAAGA
TGGCTATATGGCCTTATTGGAGCAGCGTTGCCGCAGCGCTGCACAGGCCTTTACAGAGTTGCTGAACGGT
TTAGATCCTCAAAAAATAAAGCAATTGAACCTGGCCATGATTAACTATGTTTTGGTCGTCTATGGACTTG
CCATTTCTCTCCTTGGAATAGGACAGCCTGAGGAATTATCTGAAGCCGAAAACCAGTTTAAGAGGATTAT
TGAACACTACCCCAGTGAGGGCCTTGATTGCTTGGCCTACTGTGGAATTGGAAAAGTGTATTTGAAAAAA
AACAGATTTCTAGAAGCTCTCAATCACTTTGAGAAAGCAAGAACCTTGATTTATCGTCTTCCTGGAGTGT
TAACTTGGCCCACGAGTAATGTGATTATTGAAGAGTCTCAGCCACAAAAAATAAAGATGCTGTTAGAGAA
ATTTGTTGAAGAATGCAAGTTCCCTCCAGTGCCAGATGCCATTTGTTGCTATCAGAAGTGCCATGGATAT
TCTAAGATCCAGATATACATAACTGATCCAGACTTTAAGGGTTTTATACGCATCAGCTGTTGCCAGTACT
GTAAAATAGAATTTCACATGAATTGCTGGAAGAAGTTAAAAACTACAACCTTTAATGATAAAATTGACAA
GGATTTTCTACAAGGAATATGTCTTACCCCTGACTGTGAAGGTGTCATTTCTAAGATTATCATCTTCAGC
AGTGGTGGTGAAGTTAAATGTGAATTTGAACACAAGGTCATAAAAGAAAAGGTTCCTCCAAGACCTATTC
TGAAACAGAAATGTTCTAGCCTAGAGAAACTAAGACTGAAAGAAGACAAAAAATTGAAGAGAAAGATCCA
AAAAAAAGAAGCAAAAAAGTTAGCACAAGAAAGAATGGAGGAGGACTTAAGAGAAAGTAATCCACCCAAA
AATGAAGAGCAGAAAGAAACTGTAGACAATGTTCAGCGTTGTCAGTTCCTTGATGACAGAATTCTACAGT
GTATAAAGCAGTATGCTGACAAGATTAAATCCGGCATACAGAATACAGCCATGCTTCTCAAAGAATTGCT
TTCTTGGAAAGTTTTGAGCACAGAAGACTATACAACCTGTTTTTCTAGCAGAAATTTTCTAAATGAAGCA
GTGGACTATGTTATTCGCCACTTGATTCAAGAAAATAACAGAGTAAAGACAAGAATATTTCTGCATGTTT
TGAGTGAGCTTAAAGAAGTGGAGCCCAAATTAGCCGCCTGGATCCAAAAACTTAATAGCTTTGGCTTAGA
TGCCACAGGAACTTTCTTTTCTCGTTATGGAGCATCTCTTAAACTGCTTGATTTTAGTATCATGACTTTC
CTCTGGAATGAGAAATATGGTCACAAACTAGACTCTATAGAAGGAAAGCAACTTGATTATTTCTCTGAGC
CAGCATCATTGAAGGAAGCCCGTTGTTTAATATGGCTGCTAGAAGAACACAGAGACAAGTTCCCAGCATT
GCATAGTGCTTTAGATGAATTCTTTGATATAATGGACAGCCGCTGTACTGTGTTAAGGAAACAAGATAGT
GGTGAAGCACCGTTTAGTTCAACCAAGGTGAAAAACAAAAGCAAGAAAAAGAAGCCAAAGGATTCAAAGC
CTATGTTAGTTGGGTCTGGAACAACTTCAGTAACTTCAAATAATGAGATCATCACTTCAAGTGAAGACCA
TAGCAATCGAAATTCAGATTCTGCAGGCCCATTTGCAGTGCCTGACCATCTTCGGCAAGATGTAGAAGAA
TTCGAAGCTCTCTATGACCAACACAGTAACGAATATGTTGTCCGCAATAAGAAGCTATGGGACATGAACC
CAAAACAAAAATGTTCAACTCTATATGATTACTTCTCTCAGTTTTTGGAGGAACATGGTCCCTTGGACAT
GAGTAACAAGATGTTCTCTGCAGAATATGAGTTTTTCCCAGAAGAAACTCGACAGATACTAGAAAAAGCA
GGAGGTTTAAAACCTTTTCTCTTGGGATGCCCTCGTTTTGTTGTGATTGACAACTGTATTGCACTGAAGA
AGGTTGCATCACGGCTCAAGAAAAAAAGGAAGAAGAAAAACATTAAAACAAAAGTAGAAGAAATTTCAAA
AGCAGGGGAGTATGTACGAGTTAAACTACAACTGAATCCAGCTGCTAGGGAATTTAAACCAGATGTAAAG
TCTAAACCAGTGTCAGATTCATCTTCAGCACCAGCTTTTGAAAATGTGAAACCCAAACCTGTGTCTGCAA
ATTCTCCCAAGCCAGCTTGTGAAGATGTGAAGGCCAAACCAGTATCCGACAATTCTTCTAGACAAGTTTC
TGAGGATGGGCAACCCAAAGGGGTCTCTTCTAATTCTCCTAAACCAGGCTCTGAGGATGCAAATTACAAG
CGAGTCTCCTGTAATTCCCCCAAACCGGTTCTTGAGGATGTGAAACCAACTTATTGGGCTCAATCCCATT
TGGTCACAGGATACTGTACGTATCTTCCTTTCCAGAGATTTGATATCACCCAGACACCGCCAGCATACAT
AAACGTGTTACCAGGTTTGCCCCAGTACACCAGCATATATACACCCTTGGCCAGCCTTTCTCCTGAATAT
CAGCTACCAAGATCAGTACCAGTGGTGCCGTCTTTTGTAGCCAATGACAGAGCAGATAAAAATGCTGCTG
CCTATTTTGAGGGTCATCATTTGAATGCTGAGAATGTTGCTGGTCACCAGATTGCCTCTGAAACACAGAT
CCTTGAGGGCTCTTTGGGAATATCTGTAAAGTCACACTGCAGCACAGGTGATGCTCATACAGTCCTGAGT
GAGTCTAACAGAAATGATGAGCACTGTGGAAATTCTAACAACAAATGTGAAGTAATTCCAGAAAGCACCA
GTGCAGTAACAAACATTCCACACGTGCAGATGGTTGCCATACAGGTATCTTGGAACATAATACACCAAGA
AGTCAATACTGAGCCATATAATCCTTTTGAGGAACGACAAGGGGAAATTTCACGGATTGAAAAGGAGCAC
CAAGTATTACAAGACCAACTTCAAGAAGTGTATGAAAATTATGAGCAGATAAAACTTAAGGGCTTAGAAG
AGACCAGGGACCTGGAAGAGAAGTTGAAAAGGCACTTAGAAGAAAACAAGATCTCAAAGACGGAATTAGA
TTGGTTCCTTCAAGATTTGGAAAGAGAAATTAAAAAATGGCAACAGGAAAAAAAAGAAATCCAAGAAAGA
CTAAAATCACTGAAGAAGAAAATTAAAAAGGTTTCAAATGCCAGTGAAATGTATACCCAGAAAAATGATG
GAAAGGAAAAGGAACATGAATTACATCTGGATCAGTCCCTTGAAATCAGCAACACACTTACAAATGAGAA

AATGAAAATAGAAGAGTATATAAAGAAAGGGAAAGAGGATTATGAAGAGAGTCATCAGAGAGCTGTGGCT
GCAGAGGTATCCGTACTTGAAAACTGGAAGGAGAGTGAAGTGTATAAGCTACAGATCATGGAGTCACAAG
CAGAAGCCTTTCTGAAGAAGCTGGGGCTGATTAGCCGTGATCCTGCAGCATATCCTGACATGGAGTCTGA
TATACGTTCATGGGAATTGTTTCTTTCTAATGTTACAAAAGAAATTGAGAAAGCAAAGTCTCAGTTTGAA
GAACAAATTAAGGCAATTAAAAATGGTTCTCGGCTCAGTGAACTTTCTAAAGTGCAGATTTCTGAGCTTT
CATTTCCTGCCTGTAACACGGTTCATCCCGAGTTACTCCCTGAGTCTTCAGGCCACGATGGCCAAGGGCT
TGTGACTTCTGCAAGCGACGTGACTGGAAACCACGCAGCACTTCACAGGGATCCTAGTGTGTTCTCTGCT
GGTGATTCCCCAGGGGAGGCTCCTTCTGCGCTGTTGCCAGGGCCACCCCCTGGTCAGCCTGAAGCCACTC
AGCTGACAGGGCCAAAACGGGCTGGCCAGGCAGCTCTGTCAGAACGAAGCCCTGTGGCTGATCGGAAGCA
GCCTGTTCCTCCAGGACGTGCTGCGCGTTCAAGCCAGTCTCCAAAAAAGCCGTTCAATAGTATTATTGAG
CACCTGTCAGTGGTATTCCCATGTTACAACAGCACTGAGCTTGCTGGTTTTATTAAAAAAGTGCGAAGCA
AAAACAAGAACTCACTCTCAGGATTGAGTATTGATGAAATTGTCCAAAGAGTGACAGAACACATTCTAGA
TGAACAGAAAAAGAAAAAGCCAAACCCAGGAAAGGACAAGAGGACTTATGAGCCCAGCTCTGCCACCCCC
GTGACCAGGTCCTCCCAGGGCTCACCCTCGGTGGTTGTTGCACCATCACCCAAAACCAAGGGGCAGAAAG
CAGAAGATGTCCCTGTGAGGATTGCACTGGGTGCAAGTTCCTGTGAAATATGCCACGAGGTGTTCAAATC
AAAAAACGTGCGTGTGCTCAAATGTGGGCACAAGTATCACAAAGGGTGCTTTAAGCAGTGGCTTAAAGGG
CAGAGCGCTTGCCCGGCCTGCCAGGGTCGTGATCTCCTGACAGAAGAGTCACCTTCTGGAAGAGGCTGGC
CCAGTCAGAATCAGGAGCTGCCTTCCTGCTCTTCTAGGTAGTCACACTTCACTAAAGTGTCATCCACCAG
TGTGTTGAATCCGAAGAATGACAATTTTCTACCACTGGTGTAAAAAACAAACATTTGAAGACCCTTGTGC
ATTGTGTGTCACAAAGCTAAATACATGGAAATCGTTAATATCGCTGATATTAAGTAATTTCCCCACTCTG
AGTGAATACTTTGATGATTGCCAACAGTGGCTAATAAAATGACGGCTACCACACTCATGGGTCACTGGGG
CTGCGCAGGGCTCTTTGAGGTGGGTGGCTTCTTTTGGAAAGTACTATGAACGTCTCGAAGCAGTATTCTA
GTGATAAGAATTCTTAACATAGCCAAGCGCCCCACGTTTGTTCCCCACGTTTGTTCCCCTTTTCTGTTTG
AAAAACCTGTTCTGGTAGCTCCACAAGAGAGATGATACTGACTTTTTAAATTTTTACAAGAGTCTGTAT
TCCTGATATGCCTATATTTTTCCTCAAAGATTCTGCATTTTAAGGATGGGCATAAGCAAACTATATTTTA
ATAATTTATAGTTAATGTTAAAATATTGGCTGATTTAGACCAAAAGATTCAAATCTCCTCTTTGTGAAAT
CCCATCTGCATTTGATTTTTTATTATTTTATGTTCCCCCGTTAGATTGTTTTAAGTGTTTGCTTTTCATC
TTTTATAGATGTAATCTGATTTTCAAAAATCATTAACACTTTTTAATTAGTATCGACTAAGACTTTTTCC
CCCTGGAATCGAGGCTGTGTGTCCGTCATCCCAGCCCCCGGTTGGAGCCTGCTCTTTGAACTCCGCTGCC
TTCCTTAGCAGCTTCTGTCCTCTTCTGTGAGTCAGTCAGCGAGTGCTTGGGATCCGCATCCAGCCGTGCT
GAGCACACAACAGGCTGTGTGTGGAAATGGCCACCACCATTCTCCTTCCCCACCCCACCACAAAAAGAGA
AGCTGTGTCTTTAGACAACCCTGAGGTATCTGTGTTACAATCGTTCTGTGTTTGATATTTGTGTAAAGTA
TGCATGCAGTCTTGTACTGTGACCTAAGAACAAAACTGTAACTGCATTAGAAACCATGAAAAAATTAGAT
ATTGTTTTGTGACTTTTAGACAGTGGTAAATATAGAACCATGAATTCTGGTCACATTCCATTTCTCTCCA
ACATGAAGGATCAAAAAATGTTTTTCAATGTGTTCTTTGTTCCACTGGAAACTTAGAGTCATGAGTTTAT
GAGCTGATTTGGTCACCTTCCTCTGCCTTTGTTCACTGTGAGTTCTGATGTCTTAGTGACTTAGTTCTTA
GAAGCTCACGCCTTAGTTTGAAACAGATTCTCCACGGTGGTCCCCAAAACACTGTCTGCATATCCATAAG
AATTGAGCGCTATGGGTGTTAACGTGCATGAGGATCAGTTTGCAGCAGCAAGTACAAAAGGAGAAGAGGA
ACATCCGTTGAATGAGTGTGTTTTGTACATAACTTCAGATACTTGTGAACATGCCTTATATTTGTCCAAC
AACTGTCAGAATAAAGAACATTCTAAAATGAG

Human TTC3 mRNA sequence - var4 (public gi: 1632761) (SEQ ID NO: 205)

CTGAACTAGTTGCCAGTGATCTTGAAACGTGACAGTAACCAAGAGATAAATAGGTGACAATGACAGGAAA
ATTAGATGTAGTAAAAGAGAGTGTTTGAGAGCAGAAGCTATGGCAACTAAAGACTGGATTTGAATCCTTC
CTAGCTTGGTGACATGAGCAAATTACTTGATTTAAGTGAGCATTTTCCCATCTGTACAGTGGAGATAACG
ATAATTGTGCCTGCTAAGAAGAATTGCTGTGAAGATTAGTGAAATAATGCATGTAAAACATTTGGTACAG
TATGTGACACATAGTACAAATAGTTTGCTAGGAAGATTGTTATTATTCTTCACTTGTGATATTGTGAAGT
TTTCATACAGCAAATTGGACATCATGAGATGGATTGATTAAATAAATAGATTTGAACTTCAAGGACTGGT
AGTGTTCTTGCTTTGGAAAGAAGAAACTTGGTTTATCCTAATAATAGTAGGATAATAATGGTGAAGTGAT
AGGTACAAGTAATAGTGTTTATGATGCGCTGGTGATGATAGGAAAAGAAAGCCATTATATGGGCAAGAGC
TAGAAGTAATAAAATGGTGCATTTTTCAGTGATGTTTGGCCTATGTAGCTATTCTCTGATAACTATAAAA
ATCCTTATTATTGAAGATTCTTCAGGAAAAAAAAACCCTTAGTCTGAAACTTTAGCACCAATCCCCCTTG
CCCCCCATTGAAATACGTATTTTTAAAACATGGCTTTTGATAATGTGAGGGTTTTTTCCTTTTTGCGATT
TAGCAGTGCTGATTGTGTATTGCAGTAGTTGTGAGAGCATTAGAAGCAGCAGTCGATAGGAGGATGGAAG
GTCTGGATGCCGCCTTGGGGAGTTAGGAGATTGGCAGACTTACCCTGTACCACTCTAGCCCTACTCCTTT
GCCCAAGACAGAAACACACTGAGATGGATAGGAGAATATGAGCAGTTGATAGGAAAGTTCTCAGTGGAGT
CAGGATTTAGGTTAGGCCAGGAGATTGAGAATATAACAGTTTGTGTATGATGAAATGGCATATTTCACAG
AATGCAGTAAAAGCAGGTAGGGTACAAGTGCAGCAACAGGAAGATGTCTTTTCTTCATTCAGCAAACACT
TATTTGAGAGCTTACCATGTGCTAGGCACATACAAAGATAAATAAGATGCCCTTGATGATCCTCTATTTA
AAGGAGACATGTAAACAGGTTAACTTAGAGTAGAGATGGTGAATATGTGAACCTGAGGAAAGGAAGAAAT
AGATTAAATTATCTGGAGAGAGAGGAAAAGTCAGCAGAATGGGGACGAGAATCTTTCGGAGCTCAGTGTT
CTGATAGGAGTTATTTCCTTGGGCATAGGTTCCAAGTATTTTTCTAATATACCATAGAAGCCAGGAAAAC
TTTCTTCTGTTATCTCAAATGATTTAATTACTGACTTGAGTTTGTGTTGTCTCCTTAGACTTGTGCACCA

TGGACAATTTTGCTGAGGGAGATTTCACTGTGGCGGATTATGCCTTGTTAGAAGATTGCCCTCACGTGGA
TGATTGTGTCTTTGCTGCTGAATTTATGAGCAATGATTATGTTCGTGTGACTCAGCTTTACTGTGATGGG
GTGGGTGTGCAATATAAAGATTATATCCAAAGTGAGAGGAATTTGGAATTTGACATCTGCAGTATATGGT
GTAGTAAACCAATTTCTGTCCTGCAAGATTATTGCGATGCCATTAAAATAAACATCTTCTGGCCACTTCT
GTTTCAACATCAAAACAGTTCCGTAATATCACGATTGCATCCCTGTGTGGACGCCAACAATTCACGTGCT
TCTGAGATAAATTTGAAGAAACTACAACATCTTGAGTTGATGGAAGATATTGTGGATTTGGCAAAGAAAG
TTGCTAATGATTCATTCCTTATTGGAGGCTTATTGAGAATTGGTTGTAAAATAGAAAATAAAATCTTGGC
AATGGAAGAAGCTCTGAATTGGATAAAATATGCAGGCGATGTAACAATTCTAACTAAATTAGGATCAATT
GACAATTGTTGGCCTATGTTAAGTATTTTCTTTACTGAATACAAGTACCACATAACTAAAATTGTAATGG
AAGACTGCAATTTGCTTGAAGAACTTAAAACTCAAAGTTGTATGGATTGTATAGAGGAAGGAGAACTAAT
GAAAATGAAAGGAAATGAAGAGTTTTCCAAAGAAAGATTTGATATAGCTATTATCTATTACACCAGAGCC
ATTGAATATAGACCTGAAAACTACCTTCTTTATGGTAACCGAGCTCTTTGTTTTCTTCGTACTGGACAGT
TTAGAAATGCACTCGGTGATGGAAAGAGAGCCACTATTCTGAAGAACACTTGGCCAAAGGGTCATTATCG
TTATTGTGATGCTCTTTCTATGCTGGGGGAATATGACTGGGCCCTGCAAGCAAACATAAAAGCTCAAAAA
CTCTGTAAAAATGACCCTGAGGGAATCAAGGATCTAATTCAGCAGCATGTAAAGTTACAAAAACAAATAG
AAGACCTACAAGGTCGAACAGCAAATAAGGATCCAATTAAAGCCTTTTATGAAAACAGGGCCTACACACC
TAGGAGTTTATCAGCACCTATATTTACTACTTCACTTAACTTTGTGGAGAAGGAAAGAGATTTCAGAAAA
ATTAATCACGAAATGGCCAACGGTGGTAATCAGAATCTAAAGGTGGCGGATGAGGCGTTGAAGGTAGATG
ATTGTGACTGTCATCCTGAATTTTCACCACCATCAAGTCAGCCTCCAAAACATAAAGGAAAACAAAAATC
TCGAAACAATGAATCAGAAAAGTTCAGTTCTAGTTCACCATTGACTTTACCAGCAGATTTGAAGAACATC
TTGGAGAAACAGTTTTCTAAATCTTCCAGAGCTGCACACCAGGATTTTGCTAATATAATGAAAATGCTGA
GAAGCTTAATTCAAGATGGCTATATGGCCTTATTGGAGCAGCGTTGCCGCAGCGCTGCACAGGCCTTTAC
AGAGTTGCTGAACGGTTTAGATCCTCAAAAAATAAAGCAATTGAACCTGGCCATGATTAACTATGTTTTG
GTCGTCTATGGACTTGCCATTTCTCTCCTTGGAATAGGACAGCCTGAGGAATTATCTGAAGCCGAAAACC
AGTTTAAGAGGATTATTGAACACTACCCCAGTGAGGGCCTTGATTGCTTGGCCTACTGTGGAATTGGAAA
AGTGTATTTGAAAAAAAACAGATTTCTAGAAGCTCTCAATCACTTTGAGAAAGCAAGAACCTTGATTTAT
CGTCTTCCTGGAGTGTTAACTTGGCCCACGAGTAATGTGATTATTGAAGAGTCTCAGCCACAAAAAATAA
AGATGCTGTTAGAGAAATTTGTTGAAGAATGCAAGTTCCCTCCAGTGCCAGATGCCATTTGTTGCTATCA
GAAGTGCCATGGATATTCTAAGATCCAGATATACATAACTGATCCAGACTTTAAGGGTTTTATACGCATC
AGCTGTTGCCAGTACTGTAAAATAGAATTTCACATGAATTGCTGGAAGAAGTTAAAAACTACAACCTTTA
ATGATAAAATTGACAAGGATTTTCTACAAGGAATATGTCTTACCCCTGACTGTGAAGGTGTCATTTCTAA
GATTATCATCTTCAGCAGTGGTGGTGAAGTTAAATGTGAATTTGAACACAAGGTCATAAAAGAAAAGGTT
CCTCCAAGACCTATTCTGAAACAGAAATGTTCTAGCCTAGAGAAACTAAGACTGAAAGAAGACAAAAAAT
TGAAGAGAAAGATCCAAAAAAAAGAAGCAAAAAAGTTAGCACAAGAAAGAATGGAGGAGGACTTAAGAGA
AAGTAATCCACCCAAAAATGAAGAGCAGAAAGAAACTGTAGACAATGTTCAGCGTTGTCAGTTCCTTGAT
GACAGAATTCTACAGTGTATAAAGCAGTATGCTGACAAGATTAAATCCGGCATACAGAATACAGCCATGC
TTCTCAAAGAATTGCTTTCTTGGAAAGTTTTGAGCACAGAAGACTATACAACCTGTTTTTCTAGCAGAAA
TTTTCTAAATGAAGCAGTGGACTATGTTATTCGCCACTTGATTCAAGAAAATAACAGAGTAAAGACAAGA
ATATTTCTGCATGTTTTGAGTGAGCTTAAAGAAGTGGAGCCCAAATTAGCCGCCTGGATCCAAAAACTTA
ATAGCTTTGGCTTAGATGCCACAGGAACTTTCTTTTCTCGTTATGGAGCATCTCTTAAACTGCTTGATTT
TAGTATCATGACTTTCCTCTGGAATGAGAAATATGGTCACAAACTAGACTCTATAGAAGGAAAGCAACTT
GATTATTTCTCTGAGCCAGCATCATTGAAGGAAGCCCGTTGTTTAATATGGCTGCTAGAAGAACACAGAG
ACAAGTTCCCAGCATTGCATAGTGCTTTAGATGAATTCTTTGATATAATGGACAGCCGCTGTACTGTGTT
AAGGAAACAAGATAGTGGTGAAGCACCGTTTAGTTCAACCAAGGTGAAAAACAAAAGCAAGAAAAAGAAG
CCAAAGGATTCAAAGCCTATGTTAGTTGGGTCTGGAACAACTTCAGTAACTTCAAATAATGAGATCATCA
CTTCAAGTGAAGACCATAGCAATCGAAATTCAGATTCTGCAGGCCCATTTGCAGTGCCTGACCATCTTCG
GCAAGATGTAGAAGAATTCGAAGCTCTCTATGACCAACACAGTAACGAATATGTTGTCCGCAATAAGAAG
CTATGGGACATGAACCCAAAACAAAAATGTTCAACTCTATATGATTACTTCTCTCAGTTTTTGGAGGAAC
ATGGTCCCTTGGACATGAGTAACAAGATGTTCTCTGCAGAATATGAGTTTTTCCCAGAAGAAACTCGACA
GATACTAGAAAAAGCAGGAGGTTTAAAACCTTTTCTCTTGGGATGCCCTCGTTTTGTTGTGATTGACAAC
TGTATTGCACTGAAGAAGGTTGCATCACGGCTCAAGAAAAAAAGGAAGAAGAAAAACATTAAAACAAAAG
TAGAAGAAATTTCAAAAGCAGGGGAGTATGTACGAGTTAAACTACAACTGAATCCAGCTGCTAGGGAATT
TAAACCAGATGTAAAGTCTAAACCAGTGTCAGATTCATCTTCAGCACCAGCTTTTGAAAATGTGAAACCC
AAACCTGTGTCTGCAAATTCTCCCAAGCCAGCTTGTGAAGATGTGAAGGCCAAACCAGTATCCGACAATT
CTTCTAGACAAGTTTCTGAGGATGGGCAACCCAAAGGGGTCTCTTCTAATTCTCCTAAACCAGGCTCTGA
GGATGCAAATTACAAGCGAGTCTCCTGTAATTCCCCCAAACCGGTTCTTGAGGATGTGAAACCAACTTAT
TGGGCTCAATCCCATTTGGTCACAGGATACTGTACGTATCTTCCTTTCCAGAGATTTGATATCACCCAGA
CACCGCCAGCATACATAAACGTGTTACCAGGTTTGCCCCAGTACACCAGCATATATACACCCTTGGCCAG
CCTTTCTCCTGAATATCAGCTACCAAGATCAGTACCAGTGGTGCCGTCTTTTGTAGCCAATGACAGAGCA
GATAAAAATGCTGCTGCCTATTTTGAGGGTCATCATTTGAATGCTGAGAATGTTGCTGGTCACCAGATTG
CCTCTGAAACACAGATCCTTGAGGGCTCTTTGGGAATATCTGTAAAGTCACACTGCAGCACAGGTGATGC
TCATACAGTCCTGAGTGAGTCTAACAGAAATGATGAGCACTGTGGAAATTCTAACAACAAATGTGAAGTA
ATTCCAGAAAGCACCAGTGCAGTAACAAACATTCCACACGTGCAGATGGTTGCCATACAGGTATCTTGGA

ACATAATACACCAAGAAGTCAATACTGAGCCATATAATCCTTTTGAGGAACGACAAGGGGAAATTTCACG
GATTGAAAAGGAGCACCAAGTATTACAAGACCAACTTCAAGAAGTGTATGAAAATTATGAGCAGATAAAA
CTTAAGGGCTTAGAAGAGACCAGGGACCTGGAAGAGAAGTTGAAAAGGCACTTAGAAGAAAACAAGATCT
CAAAGACGGAATTAGATTGGTTCCTTCAAGATTTGGAAAGAGAAATTAAAAAATGGCAACAGGAAAAAAA
AGAAATCCAAGAAAGACTAAAATCACTGAAGAAGAAAATTAAAAAGGTTTCAAATGCCAGTGAAATGTAT
ACCCAGAAAAATGATGGAAAGGAAAAGGAACATGAATTACATCTGGATCAGTCCCTTGAAATCAGCAACA
CACTTACAAATGAGAAAATGAAAATAGAAGAGTATATAAAGAAAGGGAAAGAGGATTATGAAGAGAGTCA
TCAGAGAGCTGTGGCTGCAGAGGTATCCGTACTTGAAAACTGGAAGGAGAGTGAAGTGTATAAGCTACAG
ATCATGGAGTCACAAGCAGAAGCCTTTCTGAAGAAGCTGGGGCTGATTAGCCGTGATCCTGCAGCATATC
CTGACATGGAGTCTGATATACGTTCATGGGAATTGTTTCTTTCTAATGTTACAAAAGAAATTGAGAAAGC
AAAGTCTCAGTTTGAAGAACAAATTAAGGCAATTAAAAATGGTTCTCGGCTCAGTGAACTTTCTAAAGTG
CAGATTTCTGAGCTTTCATTTCCTGCCTGTAACACGGTTCATCCCGAGTTACTCCCTGAGTCTTCAGGCC
ACGATGGCCAAGGGCTTGTGACTTCTGCAAGCGACGTGACTGGAAACCACGCAGCACTTCACAGGGATCC
TAGTGTGTTCTCTGCTGGTGATTCCCCAGGGGAGGCTCCTTCTGCGCTGTTGCCAGGGCCACCCCCCTGGT
CAGCCTGAAGCCACTCAGCTGACAGGGCCAAAACGGGCTGGCCAGGCAGCTCTGTCAGAACGAAGCCCTG
TGGCTGATCGGAAGCAGCCTGTTCCTCCAGGACGTGCTGCGCGTTCAAGCCAGTCTCCAAAAAAGCCGTT
CAATAGTATTATTGAGCACCTGTCAGTGGTATTCCCATGTTACAACAGCACTGAGCTTGCTGGTTTTATT
AAAAAAGTGCGAAGCAAAAACAAGAACTCACTCTCAGGATTGAGTATTGATGAAATTGTCCAAAGAGTGA
CAGAACACATTCTAGATGAACAGAAAAAGAAAAAGCCAAACCCAGGAAAGGACAAGAGGACTTATGAGCC
CAGCTCTGCCACCCCCGTGACCAGGTCCTCCCAGGGCTCACCCTCGGTGGTTGTTGCACCATCACCCAAA
ACCAAGGGGCAGAAAGCAGAAGATGTCCCTGTGAGGATTGCACTGGGTGCAAGTTCCTGTGAAATATGCC
ACGAGGTGTTCAAATCAAAAAACGTGCGTGTGCTCAAATGTGGGCACAAGTATCACAAAGGGTGCTTTAA
GCAGTGGCTTAAAGGGCAGAGCGCTTGCCCGGCCTGCCAGGGTCGTGATCTCCTGACAGAAGAGTCACCT
TCTGGAAGAGGCTGGCCCAGTCAGAATCAGGAGCTGCCTTCCTGCTCTTCTAGGTAGTCACACTTCACTA
AAGTGTCATCCACCAGTGTGTTGAATCCGAAGAATGACAATTTTCTACCACTGGTGTAAAAAACAAACAT
TTGAAGACCCTTGTGCATTGTGTGTCACAAAGCTAAATACATGGAAATCGTTAATATCGCTGATATTAAG
TAATTTCCCCACTCTGAGTGAATACTTTGATGATTGCCAACAGTGGCTAATAAAATGACGGCTACCACAC
TCATGGGTCACTGGGGCTGCGCAGGGCTCTTTGAGGTGGGTGGCTTCTTTTGGAAAGTACTATGAACGTC
TCGAAGCAGTATTCTAGTGATAAGAATTCTTAACATAGCCAAGCGCCCCACGTTTGTTCCCCACGTTTGT
TCCCCTTTTCTGTTTGAAAAACCTGTTCTGGTAGCTCCACAAGAGAGATGATACTGACTTTTTAAATTTT
TTACAAGAGTCTGTATTCCTGATATGCCTATATTTTTCCTCAAAGATTCTGCATTTTAAGGATGGGCATA
AGCAAACTATATTTTAATAATTTATAGTTAATGTTAAAATATTGGCTGATTTAGACCAAAAGATTCAAAT
CTCCTCTTTGTGAAATCCCATCTGCATTTGATTTTTTATTATTTTATGTTCCCCCGTTAGATTGTTTTAA
GTGTTTGCTTTTCATCTTTTATAGATGTAATCTGATTTTCAAAAATCATTAACACTTTTTAATTAGTATC
GACTAAGACTTTTTCCCCCTGGAATCGAGGCTGTGTGTCCGTCATCCCAGCCCCCGGTTGGAGCCTGCTC
TTTGAACTCCGCTGCCTTCCTTAGCAGCTTCTGTCCTCTTCTGTGAGTCAGTCAGCGAGTGCTTGGGATC
CGCATCCAGCCGTGCTGAGCACACAACAGGCTGTGTGTGGAAATGGCCACCACCATTCTCCTTCCCCACC
CCACCACAAAAAGAGAAGCTGTGTCTTTAGACAACCCTGAGGTATCTGTGTTACAATCGTTCTGTGTTTG
ATATTTGTGTAAAGTATGCATGCAGTCTTGTACTGTGACCTAAGAACAAAACTGTAACTGCATTAGAAAC
CATGAAAAAATTAGATATTGTTTTGTGACTTTTAGACAGTGGTAAATATAGAACCATGAATTCTGGTCAC
ATTCCATTTCTCTCCAACATGAAGGATCAAAAAATGTTTTTCAATGTGTTCTTTGTTCCACTGGAAACTT
AGAGTCATGAGTTTATGAGCTGATTTGGTCACCTTCCTCTGCCTTTGTTCACTGTGAGTTCTGATGTCTT
AGTGACTTAGTTCTTAGAAGCTCACGCCTTAGTTTGAAACAGATTCTCCACGGTGGTCCCCAAAACACTG
TCTGCATATCCATAAGAATTGAGCGCTATGGGTGTTAACGTGCATGAGGATCAGTTTGCAGCAGCAAGTA
CAAAAGGAGAAGAGGAACATCCGTTGAATGAGTGTGTTTTGTACATAACTTCAGATACTTGTGAACATGC
CTTATATTTGTCCAACAACTGTCAGAATAAAGAACATTCTAAAATGAG

Human TTC3 mRNA sequence - var5 (public gi: 2969902) (SEQ ID NO: 206)
ATATAATGTGAGGGTTTTTTCCTTTTTGCGATTTAGCAGTGCTGATTGTGTATTGCAGTAGTTGTGAGAG
CATTAGAAGCAGCAGTCGATAGGAGGATGGAAGGTCTGGATGCCGCCTTGGGGAGTTAGGAGATTGGCAG
ACTTACCCTGTACCACTCTAGCCCTACTCCTTTGCCCAAGACAGAAACACACTGAGATGGATAGGAGAAT
GTGAGCAGTTGATAGGAAAGTTCTCAGTGGAGTCAGGATTTAGGTTAGGCCAGGAGATTGAGAATATAAC
AGTTTGTGTATGATGAAATGGCATATTTCACAGAATGCAGTAAAAGCAGTGTAGGGTAAACCAAGTGCAG
TCAACAGCAAGATGTATTTTCGATGCCAGTTCAACATAAACATCTTATTGTGAGCAGTCTTACCATGTGC
TAGGCAACTATACAAAACAGATAAGATAAGATGCACGATTGACGATCCTCTATGTAAAGGACGACATGTA
CAATTCACGTGCTTAACTGAGAGTAGAGATTTGAAGAAACTACAACATCTTGAGTTGATGGAAGATATTG
TGGATTTGGCAAGGAAAGTTGCTAATGATTCATTCCTTATTGGAGGCTTATTGAGAATTGGTTGTAAAAT
AGAAAATAAAATCTTGGCAATGGAAGAAGCTCTGAATTGGATAAAATATGCAGGCGATGTAACAATTCTA
ACTAAATTAGGATCAATTGACAATTGTTGGCCTATGTTAAGTATTTTCTTTACTGAATACAAGTACCACA
TAACTAAAATTGTAATGGAAGACTGCAATTTGCTTGAAGAACTTAAAACTCAAAGTTGTATGGATTGTAT
AGAGGAAGGAGGACTAATGAAAATGAAAGGAAATGAAGAGTTTTCCAAAGAAAGATTTGATATAGCTATT
ATCTATTACACCAGAGCCATTGAATATAGACCTGAAAACTACCTTCTTTATGGTAACCGAGCTCTTTGTT
TTCCTCGTACTGGACAGTTTAGAAATGCACTCGGTGATGGAAAGAGAGCCACTATTCTGAAGAACACTTG

GCCAAAGGGTCATTATCGTTATTGTGATGCTCTTTCTATGCTGGGGGAATATGACTGGGCCCTGCAAGCA
AACATAAAAGCTCAAAAACTCTGTAAAAATGACCCTGAGGGAATCAAGGATCTAATTCAGCAGCATGTAA
AGTTACAAAAACAAATAGAAGACCTACAAGGTCGAACAGCAAATAAGGATCCAATTAAAGCCTTTTATGA
AAACAGGGCCTACACACCTAGGAGTTTATCAGCACCTATATTTACTACTTCACTTAACTTTGTGGAGAAG
GAAAGAGATTTCAGAAAAATTAATCACGAAATGGCCAACGGTGGTAATCAGAATCTAAAGGTGGCGGATG
AGGCGTTGAAGGTAGATGATTGTGACTGTCATCCTGAATTTTCACCACCATCAAGTCAGCCTCCAAAACA
TAAAGGAAAACAAAAATCTCGAAACAATGAATCAGAAAAGTTCAGTTCTAGTTCACCATTGACTTTACCA
GCAGATTTGAAGAACATCTTGGAGAAACAGTTTTCTAAATCTTCCAGAGCTGCACACCAGGATTTTGCTA
ATATAATGAAAATGCTGAGAAGCTTAATTCAAGATGGCTATATGGCCTTATTGGAGCAGCGTTGCCGCAG
CGCTGCACAGGCCTTTACAGAGTTGCTGAACGGTTTAGATCCTCAAAAAATAAAGCAATTGAACCTGGCC
ATGATTAACTATGTTTTGGTCGTCTATGGACTTGCCATTTCTCTCCTTGGAATAGGACAGCCTGAGGAAT
TATCTGAAGCCGAAAACCAGTTTAAGAGGATTATTGAACACTACCCCAGTGAGGGCCTTGATTGCTTGGC
CTACTGTGGAATTGGAAAAGTGTATTTGAAAAAAAACAGATTTCTAGAAGCTCTCAATCACTTTGAGAAA
GCAAGAACCTTGATTTATCGTCTTCCTGGAGTGTTAACTTGGCCA

Human TTC3 mRNA sequence - var6 (public gi: 1304131) (SEQ ID NO: 207)

CCTAAAGAAAAGTATTAAGTAAATAGCAGTACAGATGGCAAATGGATTGCACAATATATCCTCTGGATCC
ATAGTGACCCTGCAGAGATAAACCTGTGATGGTCAAACAATGTGAAAACTGCTGTCAGAGACATGGGCAG
GGTGCTCTTGTTTACAGAGAAGAGGTGCAAAAATCAACTTGATGGTAGTGGGAAGATCAGGAAATGCTTC
CTGAAATTGAGTATTAAGAACTAATAGACATTAGGTGGTTGCAGAATAAGTTTTGTTTAGGAAGGACAAG
CAGTTGGGTATGACTGGCTTCTAGGTTGTGTGTTGTGGAGTGACTGGGGATAAAAGCAGGAGCAAGATCA
CAAAAGGTCTTCTATGCTTATATTAGGGAAGTTGGACTTTATTCTCAAGCTGAAGGGAAGCTGTTGCATG
GTTTTAAGCAGTAAAGTGATATGATCAGAGTTTTAGAGGATGCCAAGATTGAAGGCAAGTCTGACCAGTT
AGGAGACTGCTTGTTAAATTAGTTCAGAGGAGAAACAGTGAAGGCAGTGGCACTGGGCATGAAGAAGTAT
ATGTGTGCTAATTTTAGATTTCTTAGGGAAGCAGAAATGACAAGAGTTAGTGGTCCATTGGACAGAAATA
TTGAAGGAGACTGGGGAGTCTAGGTTGACTCCCAGGGTTTAGGTTTGGGCAGTAAAATGACATGTAGAAC
AATTAACTGATAAACAGCATACAGAAAGAGGAAAGAACTTCATTTCATATGTTTGTTTTTGAGGAAAAGA
TGTTTGTTTTTGAACTTCCTGATTCAGAGGGGCTTGTGGGACATCTTGGTTAAGATCCTGTAGTAGTTCT
AGTAGGGTCTAGAAGTCAAGAGATACAACCCCGCCTGGAAGGATTTGGGAGTCTTCAGCATTTGGAATTT
TGGAAGCCATTGTTTACTGCAGTGCATATGAGATAATTTAAACTGGTACATAGATAAACACTTGAAAAAA
TTTTAATAGATAGGAATTTAATGTGTATGAGAAACATAACTTGCACATCTAACCTTTGATAATCATGGAC
ATTATCACTTAGGCCAAGTGAGCTCAATAAAAGGGAAATATTAACTAAAAATATATGTGATACATGGAAA
TGGCAAAAATCACGGTGAAGCTATGCAAACGATGCAAGTTCAGAAAGTGCTGTGTTAAGTCATGGGTGTG
GTGGATAAATTCACCCAAGGAGAGAGTAAGAGTGAGAAGAAAAGAGAGTTGACATTGGGTGGCGGGGAAT
GGAGAAGAAGAACCCATGAAGGAAACTGAGGAAGAGCAGCCAGACGAATAGGAGGAAAACCAGGAGAAGA
TGGTCTTTGGAGTCCAAATGAAGAGTTCTGAGGAGGAAGTGGTCCACAGCGTCAAACTTGTGCACCATGG
ACAATTTTGCTGAGGGAGATTTCACTGTGGCGGATTATGCCTTGTTAGAAGATTGCCCTCACGTGGATGA
TTGTGTCTTTGCTGCTGAATTTATGAGCAATGATTATGTTCGTGTGACTCAGCTTTACTGTGATGGGGTG
GGTGTGCAATATAAAGATTATATCCAAAGTGAGAGGAATTTGGAATTTGACATCTGCAGTATATGGTGTA
GTAAACCAATTTCTGTCCTGCAAGATTATTGCGATGCCATTAAAATAAACATCTTCTGGCCACTTCTGTT
TCAACATCAAAACAGTTCCGTAATATCACGATTGCATCCCTGTGTGGACGCCAACAATTCACGTGCTTCT
GAGATAAATTTGAAGAAACTACAACATCTTGAGTTGATGGAAGATATTGTGGATTTGGCAAAGAAAGTTG
CTAATGATTCATTCCTTATTGGAGGCTTATTGAGAATTGGTTGTAAAATAGAAAATAAAATCTTGGCAAT
GGAAGAAGCTCTGAATTGGATAAAATATGCAGGCGATGTAACAATTCTAACTAAATTAGGATCAATTGAC
AATTGTTGGCCTATGTTAAGTATTTTCTTTACTGAATACAAGTACCACATAACTAAAATTGTAATGGAAG
ACTGCAATTTGCTTGAAGAACTTAAAACTCAAAGTTGTATGGATTGTATAGAGGAAGGAGAACTAATGAA
AATGAAAGGAAATGAAGAGTTTTCCAAAGAAAGATTTGATATAGCTATTATCTATTACACCAGAGCCATT
GAATATAGACCTGAAAACTACCTTCTTTATGGTAACCGAGCTCTTTGTTTTCTTCGTACTGGACAGTTTA
GAAATGCACTCGGTGATGGAAAGAGAGCCACTATTCTGAAGAACACTTGGCCAAAGGGTCATTATCGTTA
TTGTGATGCTCTTTCTATGCTGGGGGAATATGACTGGGCCCTGCAAGCAAACATAAAAGCTCAAAAACTC
TGTAAAAATGACCCTGAGGGAATCAAGGATCTAATTCAGCAGCATGTAAAGTTACAAAAACAAATAGAAG
ACCTACAAGGTCGAACAGCAAATAAGGATCCAATTAAAGCCTTTTATGAAAACAGGGCCTACACACCTAG
GAGTTTATCAGCACCTATATTTACTACTTCACTTAACTTTGTGGAGAAGGAAAGAGATTTCAGAAAAATT
AATCACGAAATGGCCAACGGTGGTAATCAGAATCTAAAGGTGGCGGATGAGGCGTTGAAGGTAGATGATT
GTGACTGTCATCCTGAATTTTCACCACCATCAAGTCAGCCTCCAAAACATAAAGGAAAACAAAAATCTCG
AAACAATGAATCAGAAAAGTTCAGTTCTAGTTCACCATTGACTTTACCAGCAGATTTGAAGAACATCTTG
GAGAAACAGTTTTCTAAATCTTCCAGAGCTGCACACCAGGATTTTGCTAATATAATGAAAATGCTGAGAA
GCTTAATTCAAGATGGCTATATGGCCTTATTGGAGCAGCGTTGCCGCAGCGCTGCACAGGCCTTTACAGA
GTTGCTGAACGGTTTAGATCCTCAAAAAATAAAGCAATTGAACCTGGCCATGATTAACTATGTTTTGGTC
GTCTATGGACTTGCCATTTCTCTCCTTGGAATAGGACAGCCTGAGGAATTATCTGAAGCCGAAAACCAGT
TTAAGAGGATTATTGAACACTACCCCAGTGAGGGCCTTGATTGCTTGGCCTACTGTGGAATTGGAAAAGT
GTATTTGAAAAAAAACAGATTTCTAGAAGCTCTCAATCACTTTGAGAAAGCAAGAACCTTGATTTATCGT
CTTCCTGGAGTGTTAACTTGGCCCACGAGTAATGTGATTATTGAAGAGTCTCAGCCACAAAAAAATAAAGA

TGCTGTTAGAGAAATTTGTTGAAGAATGCAAGTTCCCTCCAGTGCCAGATGCCATTTGTTGCTATCAGAA
GTGCCATGGATATTCTAAGATCCAGATATACATAACTGATCCAGACTTTAAGGGTTTTATACGCATCAGC
TGTTGCCAGTACTGTAAAATAGAATTTCACATGAATTGCTGGAAGAAGTTAAAAACTACAACCTTTAATG
ATAAAATTGACAAGGATTTTCTACAAGGAATATGTCTTACCCCTGACTGTGAAGGTGTCATTTCTAAGAT
TATCATCTTCAGCAGTGGTGGTGAAGTTAAATGTGAATTTGAACACAAGGTCATAAAAGAAAAGGTTCCT
CCAAGACCTATTCTGAAACAGAAATGTTCTAGCCTAGAGAAACTAAGACTGAAAGAAGACAAAAAATTGA
AGAGAAAGATCCAAAAAAAAGAAGCAAAAAAGTTAGCACAAGAAAGAATGGAGGAGGACTTAAGAGAAAG
TAATCCACCCAAAAATGAAGAGCAGAAAGAAACTGTAGACAATGTTCAGCGTTGTCAGTTCCTTGATGAC
AGAATTCTACAGTGTATAAAGCAGTATGCTGACAAGATTAAATCCGGCATACAGAATACAGCCATGCTTC
TCAAAGAATTGCTTTCTTGGAAAGTTTTGAGCACAGAAGACTATACAACCTGTTTTTCTAGCAGAAATTT
TCTAAATGAAGCAGTGGACTATGTTATTCGCCACTTGATTCAAGAAAATAACAGAGTAAAGACAAGAATA
TTTCTGCATGTTTTGAGTGAGCTTAAAGAAGTGGAGCCCAAATTAGCCGCCTGGATCCAAAAACTTAATA
GCTTTGGCTTAGATGCCACAGGAACTTTCTTTTCTCGTTATGGAGCATCTCTTAAACTGCTTGATTTTAG
TATCATGACTTTCCTCTGGAATGAGAAATATGGTCACAAACTAGACTCTATAGAAGGAAAGCAACTTGAT
TATTTCTCTGAGCCAGCATCATTGAAGGAAGCCCGTTGTTTAATATGGCTGCTAGAAGAACACAGAGACA
AGTTCCCAGCATTGCATAGTGCTTTAGATGAATTCTTTGATATAATGGACAGCCGCTGTACTGTGTTAAG
GAAACAAGATAGTGGTGAAGCACCGTTTAGTTCAACCAAGGTGAAAAACAAAAGCAAGAAAAAGAAGCCA
AAGGATTCAAAGCCTATGTTAGTTGGGTCTGGAACAACTTCAGTAACTTCAAATAATGAGATCATCACTT
CAAGTGAAGACCATAGCAATCGAAATTCAGATTCTGCAGGCCCATTTGCAGTGCCTGACCATCTTCGGCA
AGATGTAGAAGAATTCGAAGCTCTCTATGACCAACACAGTAACGAATATGTTGTCCGCAATAAGAAGCTA
TGGGACATGAACCCAAAACAAAAATGTTCAACTCTATATGATTACTTCTCTCAGTTTTTGGAGGAACATG
GTCCCTTGGACATGAGTAACAAGATGTTCTCTGCAGAATATGAGTTTTTCCCAGAAGAAACTCGACAGAT
ACTAGAAAAAGCAGGAGGTTTAAAACCTTTTCTCTTGGGATGCCCTCGTTTTGTTGTGATTGACAACTGT
ATTGCACTGAAGAAGGTTGCATCACGGCTCAAGAAAAAAAGGAAGAAGAAAAACATTAAAACAAAAGTAG
AAGAAATTTCAAAAGCAGGGGAGTATGTACGAGTTAAACTACAACTGAATCCAGCTGCTAGGGAATTTAA
ACCAGATGTAAAGTCTAAACCAGTGTCAGATTCATCTTCAGCACCAGCTTTTGAAAATGTGAAACCCAAA
CCTGTGTCTGCAAATTCTCCCAAGCCAGCTTGTGAAGATGTGAAGGCCAAACCAGTATCCGACAATTCTT
CTAGACAAGTTTCTGAGGATGGGCAACCCAAAGGGGTCTCTTCTAATTCTCCTAAACCAGGCTCTGAGGA
TGCAAATTACAAGCGAGTCTCCTGTAATTCCCCCAAACCGGTTCTTGAGGATGTGAAACCAACTTATTGG
GCTCAATCCCATTTGGTCACAGGATACTGTACGTATCTTCCTTTCCAGAGATTTGATATCACCCAGACAC
CGCCAGCATACATAAACGTGTTACCAGGTTTGCCCCAGTACACCAGCATATATACACCCTTGGCCAGCCT
TTCTCCTGAATATCAGCTACCAAGATCAGTACCAGTGGTGCCGTCTTTTGTAGCCAATGACAGAGCAGAT
AAAAATGCTGCTGCCTATTTTGAGGGTCATCATTTGAATGCTGAGAATGTTGCTGGTCACCAGATTGCCT
CTGAAACACAGATCCTTGAGGGCTCTTTGGGAATATCTGTAAAGTCACACTGCAGCACAGGTGATGCTCA
TACAGTCCTGAGTGAGTCTAACAGAAATGATGAGCACTGTGGAAATTCTAACAACAAATGTGAAGTAATT
CCAGAAAGCACCAGTGCAGTAACAAACATTCCACACGTGCAGATGGTTGCCATACAGGTATCTTGGAACA
TAATACACCAAGAAGTCAATACTGAGCCATATAATCCTTTTGAGGAACGACAAGGGGAAATTTCACGGAT
TGAAAAGGAGCACCAAGTATTACAAGACCAACTTCAAGAAGTGTATGAAAATTATGAGCAGATAAAACTT
AAGGGCTTAGAAGAGACCAGGGACCTGGAAGAGAAGTTGAAAAGGCACTTAGAAGAAAACAAGATCTCAA
AGACGGAATTAGATTGGTTCCTTCAAGATTTGGAAAGAGAAATTAAAAAATGGCAACAGGAAAAAAAAGA
AATCCAAGAAAGACTAAAATCACTGAAGAAGAAAATTAAAAAGGTTTCAAATGCCAGTGAAATGTATACC
CAGAAAAATGATGGAAAGGAAAAGGAACATGAATTACATCTGGATCAGTCCCTTGAAATCAGCAACACAC
TTACAAATGAGAAAATGAAAATAGAAGAGTATATAAAGAAAGGGAAAGAGGATTATGAAGAGAGTCATCA
GAGAGCTGTGGCTGCAGAGGTATCCGTACTTGAAAACTGGAAGGAGAGTGAAGTGTATAAGCTACAGATC
ATGGAGTCACAAGCAGAAGCCTTTCTGAAGAAGCTGGGGCTGATTAGCCGTGATCCTGCAGCATATCCTG
ACATGGAGTCTGATATACGTTCATGGGAATTGTTTCTTTCTAATGTTACAAAAGAAATTGAGAAAGCAAA
GTCTCAGTTTGAAGAACAAATTAAGGCAATTAAAAATGGTTCTCGGCTCAGTGAACTTTCTAAAGTGCAG
ATTTCTGAGCTTTCATTTCCTGCCTGTAACACGGTTCATCCCGAGTTACTCCCTGAGTCTTCAGGCCACG
ATGGCCAAGGGCTTGTGACTTCTGCAAGCGACGTGACTGGAAACCACGCAGCACTTCACAGGGATCCTAG
TGTGTTCTCTGCTGGTGATTCCCCAGGGGAGGCTCCTTCTGCGCTGTTGCCAGGGCCACCCCCTGGTCAG
CCTGAAGCCACTCAGCTGACAGGGCCAAAACGGGCTGGCCAGGCAGCTCTGTCAGAACGAAGCCCTGTGG
CTGATCGGAAGCAGCCTGTTCCTCCAGGACGTGCTGCGCGTTCAAGCCAGTCTCCAAAAAAGCCGTTCAA
TAGTATTATTGAGCACCTGTCAGTGGTATTCCCATGTTACAACAGCACTGAGCTTGCTGGTTTTATTAAA
AAAGTGCGAAGCAAAAACAAGAACTCACTCTCAGGATTGAGTATTGATGAAATTGTCCAAAGAGTGACAG
AACACATTCTAGATGAACAGAAAAAGAAAAAGCCAAACCCAGGAAAGGACAAGAGGACTTATGAGCCCAG
CTCTGCCACCCCCGTGACCAGGTCCTCCCAGGGCTCACCCTCGGTGGTTGTTGCACCATCACCCAAAACC
AAGGGGCAGAAAGCAGAAGATGTCCCTGTGAGGATTGCACTGGGTGCAAGTTCCTGTGAAATATGCCACG
AGGTGTTCAAATCAAAAAACGTGCGTGTGCTCAAATGTGGGCACAAGTATCACAAAGGGTGCTTTAAGCA
GTGGCTTAAAGGGCAGAGCGCTTGCCCGGCCTGCCAGGGTCGTGATCTCCTGACAGAAGAGTCACCTTCT
GGAAGAGGCTGGCCCAGTCAGAATCAGGAGCTGCCTTCCTGCTCTTCTAGGTAGTCACACTTCACTAAAG
TGTCATCCACCAGTGTGTTGAATCCGAAGAATGACAATTTTCTACCACTGGTGTAAAAAACAAACATTTG
AAGACCCTTGTGCATTGTGTGTCACAAAGCTAAATACATGGAAATCGTTAATATCGCTGATATTAAGTAA
TTTCCCCACTCTGAGTGAATACTTTGATGATTGCCAACAGTGGCTAATAAAATGACGGCTACCACACTCA

TGGGTCACTGGGCTGCGCAGGGCTCTTTGAGGTGGGTGGCTTCTTTTGGAAAGTACTATGAACGTCTCGA
AGCAGTATTCTAGTGATAAGAATTCTTAACATAGCCAAGCGCCCCACGTTTGTTCCCCACGTTTGTTCCC
CTTTTCTGTTTGAAAAACCTGTTCTGGTAGCTCCACAAGAGAGATGATACTGACTTTTTAAATTTTTTAC
AAGAGTCTGTATTCCTGATATGCCTATATTTTTCCTCAAAGATTCTGCATTTTAAGGATGGGCATAAGCA
AACTATATTTTAATAATTTATAGTTAATGTTAAAATATTGGCTGATTTAGACCAAAAGATTCAAATCTCC
TCTTTGTGAAATCCCATCTGCATTTGATTTTTTATTATTTTATGTTCCCCCGTTAGATTGTTTTAAGTGT
TTGCTTTTCATCTTTTATAGATGTAATCTGATTTTCAAAAATCATTAACACTTTTTAATTAGTATCGACT
AAGACTTTTTCCCCCTGGAATCGAGGCTGTGTGTCCGTCATCCCAGCCCCCGGTTGGAGCCTGCTCTTTG
AACTCCGCTGCCTTCCTTAGCAGCTTCTGTCCTCTTCTGTGAGTCAGTCAGCGAGTGCTTGGGATCCGCA
TCCAGCCGTGCTGAGCACACAACAGGCTGTGTGTGGAAATGGCCACCACCATTCTCCTTCCCCACCCCAC
CACAAAAAGAGAAGCTGTGTCTTTAGACAACCCTGAGGTATCTGTGTTACAATCGTTCTGTGTTTGATAT
TTGTGTAAAGTATGCATGCAGTCTTGTACTGTGACCTAAGAACAAAACTGTAACTGCATTAGAAACCATG
AAAAAATTAGATATTGTTTTGTGACTTTTAGACAGTGGTAAATATAGAACCATGAATTCTGGTCACATTC
CATTTCTCTCCAACATGAAGGATCAAAAAATGTTTTTCAATGTGTTCTTTGTTCCACTGGGAAACTTAGA
GTCATGAGTTTATGAGGCTGGATTTGGGCACCTTTCCTTTGCCTTTGGTTCACTGTGAGTTCTGATGTCC
TAGTGACTTAGGTCTTAGAAGCTCACGCCTTAGTTTGAAACAGATTCTCCACGGTGGTCCCCAAAACACT
GTCTGCATATCCATAAGAATTGAACGCTATGGGTGTTAACGTGCATGAGGATCAGTTTGCAGCAGCAAGT
ACAAAAGGAGAAGAGGAACATCCGTTGAATGAGTGTGTTTTGTACATAACTTCAGATACTTGTGAACATG
CCTTATATTTGTCCAACAACTGTCAGAATAAAGAACATTCTAAAATGAG

Human TTC3 Protein sequence - var1 (public gi: 2662364) (SEQ ID NO: 308)
IKINIFWPLLFQHQNSSVISRLHPCVDANNSRASEINLKKLQHLELMEDIVDLAKKVANDSFLIGGLLRI
GCKIENKILAMEEALNWIKYAGDVTILTKLGSIDNCWPMLSIFFTEYKYHITKIVMEDCNLLEELKTQSC
MDCIEEGELMKMKGNEEFSKERFDIAIIYYTRAIEYRPENYLLYGNRALCFLRTGQFRNALGDGKRATIL
KNTWPKGHYRYCDALSMLGEYDWALQANIKAQKLCKNDPEGIKDLIQQHVKLQKQIEDLQGRTANKDPIK
AFYENRAYTPRSLSAPIFTTSLNFVEKERDFRKINHEMANGGNQNLKVADEALKVDDCDCHPEFSPPSSQ
PPKHKGKQKSRNNESEKFSSSSPLTLPADLKNILEKQFSKSSRAAHQDFANIMKMLRSLIPDGYMALLEQ
RCRSAAQAFTELLNGLDPQKIKQLNLAMINYVLVVYGLAISLLGIGQPEELSEAENQFKRIIEHYPSEGL
DCLAYCGIGKVYLKKNRFLEALNHFEKARTLIYRLPGVLTWPTSNVIIEESQPQKIKMLLEKFVEECKFP
PVPDAICCYQKCHGYSKIQIYITDPDFKGFIRISCCQYCKIEFHMNCWKKLKTTTFNDKIDKDFLQGICL
TPDCEGVISKIIIFSSGGEVKCEFEHKVIKEKVPPRPILKQKCSSLEKLRLKEDKKLKRKIQKKEAKKLA
QERMEEDLRESNPPKNEEQKETVDNVQRCQFLDDRILQCIKQYADKIKSGIQNTATLLKELLSWKVLSTE
DYTTCFSSRNFLNEAVDYVIRHLIQENNRVKTRIFLHVLSELKEVEPKLAAWIQKLNSFGLDATGTFFSR
YGASLKLLDFSIMTFLWNEKYGHKLDSIEGKQLDYFSEPASLKEARCLIWLLEEHRDKFPALHSALDEFF
DIMDSRCTVLRKQDSGEAPFSSTKVKNKSKKKKPKDSKPMLVGSGTTSVTSNNEIITSSEDHSNRNSDSA
GPFAVPDHLRQDVEEFEALYDQHSNEYVVRNKKLWDMNPKQKCSTLYDYFSQFLEEHGPLDMSNKMFSAE
YEFFPEETRQILEKAGGLKPFLLGCPRFVVIDNCIALKKVASRLKKKRKKKNIKTKVEEISKAGEYVRVK
LQLNPAAREFKPDVKSKPVSDSSSAPAFENVKPKPVSANSPKPACEDVKAKPVSDNSSRQVSEDGQPKGV
SSNSPKPGSEDANYKRVSCNSPKPVLEDVKPTYWAQSHLVTGYCTYLPFQRFDITQTPPAYINVLPGLPQ
YTSIYTPLASLSPEYQLPRSVPVVPSFVANDRADKNAAAYFEGHHLNAENVAGHQIASETQILEGSLGIS
VKSHCSTGDAHTVLSESNRNDEHCGNSNNKCEVIPESTSAVTNIPHVQMVAIQVSWNIIHQEVNTEPYNP
FEERQGEISRIEKEHQVLQDQLQEVYENYEQIKLKGLEETRDLEEKLKRHLEENKISKTELDWFLQDLER
EIKKWQQEKKEIQERLKSLKKKIKKVSNASEMYTQKNDGKEKEHELHLDQSLEISNTLTNEKMKIEEYIK
KGKEDYEESHQRAVAAEVSVLENWKESEVYKLQIMESQAEAFLKKLGLISRDPAAYPDMESDIRSWELFL
SNVTKVIEKAKSQFEEQIKAIKNGSRLSELSKVQISELSFPACNTVHPELLPESSGDDGQGLVTSASDVT
GNHAALHRDPSVFSAGDSPGEAPSALLPGPPPGQPEATQLTGPKRAGQAALSERSPVTDRKQPVPPGRAA
RSSQSPKKPFNSIIEHLSVVFPCYNSTELAGFIKKVRSKNKNSLSGLSIDEIVQRVTEHILDEQKKKKPN
PGKDKRTYEPSSATPVTRSSQGSPSVVVAPSPKTKGQKAEDVPVRIALGASSCEICHEVFKSKNVRVLKC
GHKYHKGCFKQWLKGQSACPACQGRDLLTEESPSGRGWPSQNQELPSCSSR

Human TTC3 Protein sequence - var2 (public gi: 1632766) (SEQ ID NO: 309)
MLGEYDWALQANIKAQKLCKNDPEGIKDLIQQHVKLQKQIEDLQGRTANKDPIKAFYENRAYTPRSLSAP
IFTTSLNFVEKERDFRKINHEMANGGNQNLKVADEALKVDDCDCHPEFSPPSSQPPKHKGKQKSRNNESE
KFSSSSPLTLPADLKNILEKQFSKSSRAAHQDFANIMKMLRSLIQDGYMALLEQRCRSAAQAFTELLNGL
DPQKIKQLNLAMINYVLVVYGLAISLLGIGQPEELSEAENQFKRIIEHYPSEGLDCLAYCGIGKVYLKKN
RFLEALNHFEKARTLIYRLPGVLTWPTSNVIIEESQPQKIKMLLEKFVEECKFPPVPDAICCYQKCHGYS
KIQIYITDPDFKGFIRISCCQYCKIEFHMNCWKKLKTTTFNDKIDKDFLQGICLTPDCEGVISKIIIFSS
GGEVKCEFEHKVIKEKVPPRPILKQKCSSLEKLRLKEDKKLKRKIQKKEAKKLAQERMEEDLRESNPPKN
EEQKETVDNVQRCQFLDDRILQCIKQYADKIKSGIQNTAMLLKELLSWKVLSTEDYTTCFSSRNFLNEAV
DYVIRHLIQENNRVKTRIFLHVLSELKEVEPKLAAWIQKLNSFGLDATGTFFSRYGASLKLLDFSIMTFL
WNEKYGHKLDSIEGKQLDYFSEPASLKEARCLIWLLEEHRDKFPALHSALDEFFDIMDSRCTVLRKQDSG
EAPFSSTKVKNKSKKKKPKDSKPMLVGSGTTSVTSNNEIITSSEDHSNRNSDSAGPFAVPDHLRQDVEEF

EALYDQHSNEYVVRNKKLWDMNPKQKCSTLYDYFSQFLEEHGPLDMSNKMFSAEYEFFPEETRQILEKAG
GLKPFLLGCPRFVVIDNCIALKKVASRLKKKRKKKNIKTKVEEISKAGEYVRVKLQLNPAAREFKPDVKS
KPVSDSSSAPAFENVKPKPVSANSPKPACEDVKAKPVSDNSSRQVSEDGQPKGVSSNSPKPGSEDANYKR
VSCNSPKPVLEDVKPTYWAQSHLVTGYCTYLPFQRFDITQTPPAYINVLPGLPQYTSIYTPLASLSPEYQ
LPRSVPVVPSFVANDRADKNAAAYFEGHHLNAENVAGHQIASETQILEGSLGISVKSHCSTGDAHTVLSE
SNRNDEHCGNSNNKCEVIPESTSAVTNIPHVQMVAIQVSWNIIHQEVNTEPYNPFEERQGEISRIEKEHQ
VLQDQLQEVYENYEQIKLKGLEETRDLEEKLKRHLEENKISKTELDWFLQDLEREIKKWQQEKKEIQERL
KSLKKKIKKVSNASEMYTQKNDGKEKEHELHLDQSLEISNTLTNEKMKIEEYIKKGKEDYEESHQRAVAA
EVSVLENWKESEVYKLQIMESQAEAFLKKLGLISRDPAAYPDMESDIRSWELFLSNVTKEIEKAKSQFEE
QIKAIKNGSRLSELSKVQISELSFPACNTVHPELLPESSGHDGQGLVTSASDVTGNHAALHRDPSVFSAG
DSPGEAPSALLPGPPPGQPEATQLTGPKRAGQAALSERSPVADRKQPVPPGRAARSSQSPKKPFNSIIEH
LSVVFPCYNSTELAGFIKKVRSKNKNSLSGLSIDEIVQRVTEHILDEQKKKKPNPGKDKRTYEPSSATPV
TRSSQGSPSVVVAPSPKTKGQKAEDVPVRIALGASSCEICHEVFKSKNVRVLKCGHKYHKGCFKQWLKGQ
SACPACQGRDLLTEESPSGRGWPSQNQELPSCSSR

Human TTC3 Protein sequence - var3 (public gi: 1632764) (SEQ ID NO: 310)
MKMKGNEEFSKERFDIAIIYYTRAIEYRPENYLLYGNRALCFLRTGQFRNALGDGKRATILKNTWPKGHY
RYCDALSMLGEYDWALQANIKAQKLCKNDPEGIKDLIQQHVKLQKQIEDLQGRTANKDPIKAFYENRAYT
PRSLSAPIFTTSLNFVEKERDFRKINHEMANGGNQNLKVADEALKVDDCDCHPEFSPPSSQPPKHKGKQK
SRNNESEKFSSSSPLTLPADLKNILEKQFSKSSRAAHQDFANIMKMLRSLIQDGYMALLEQRCRSAAQAF
TELLNGLDPQKIKQLNLAMINYVLVVYGLAISLLGIGQPEELSEAENQFKRIIEHYPSEGLDCLAYCGIG
KVYLKKNRFLEALNHFEKARTLIYRLPGVLTWPTSNVIIEESQPQKIKMLLEKFVEECKFPPVPDAICCY
QKCHGYSKIQIYITDPDFKGFIRISCCQYCKIEFHMNCWKKLKTTTFNDKIDKDFLQGICLTPDCEGVIS
KIIIFSSGGEVKCEFEHKVIKEKVPPRPILKQKCSSLEKLRLKEDKKLKRKIQKKEAKKLAQERMEEDLR
ESNPPKNEEQKETVDNVQRCQFLDDRILQCIKQYADKIKSGIQNTAMLLKELLSWKVLSTEDYTTCFSSR
NFLNEAVDYVIRHLIQENNRVKTRIFLHVLSELKEVEPKLAAWIQKLNSFGLDATGTFFSRYGASLKLLD
FSIMTFLWNEKYGHKLDSIEGKQLDYFSEPASLKEARCLIWLLEEHRDKFPALHSALDEFFDIMDSRCTV
LRKQDSGEAPFSSTKVKNKSKKKKPKDSKPMLVGSGTTSVTSNNEIITSSEDHSNRNSDSAGPFAVPDHL
RQDVEEFEALYDQHSNEYVVRNKKLWDMNPKQKCSTLYDYFSQFLEEHGPLDMSNKMFSAEYEFFPEETR
QILEKAGGLKPFLLGCPRFVVIDNCIALKKVASRLKKKRKKKNIKTKVEEISKAGEYVRVKLQLNPAARE
FKPDVKSKPVSDSSSAPAFENVKPKPVSANSPKPACEDVKAKPVSDNSSRQVSEDGQPKGVSSNSPKPGS
EDANYKRVSCNSPKPVLEDVKPTYWAQSHLVTGYCTYLPFQRFDITQTPPAYINVLPGLPQYTSIYTPLA
SLSPEYQLPRSVPVVPSFVANDRADKNAAAYFEGHHLNAENVAGHQIASETQILEGSLGISVKSHCSTGD
AHTVLSESNRNDEHCGNSNNKCEVIPESTSAVTNIPHVQMVAIQVSWNIIHQEVNTEPYNPFEERQGEIS
RIEKEHQVLQDQLQEVYENYEQIKLKGLEETRDLEEKLKRHLEENKISKTELDWFLQDLEREIKKWQQEK
KEIQERLKSLKKKIKKVSNASEMYTQKNDGKEKEHELHLDQSLEISNTLTNEKMKIEEYIKKGKEDYEES
HQRAVAAEVSVLENWKESEVYKLQIMESQAEAFLKKLGLISRDPAAYPDMESDIRSWELFLSNVTKEIEK
AKSQFEEQIKAIKNGSRLSELSKVQISELSFPACNTVHPELLPESSGHDGQGLVTSASDVTGNHAALHRD
PSVFSAGDSPGEAPSALLPGPPPGQPEATQLTGPKRAGQAALSERSPVADRKQPVPPGRAARSSQSPKKP
FNSIIEHLSVVFPCYNSTELAGFIKKVRSKNKNSLSGLSIDEIVQRVTEHILDEQKKKKPNPGKDKRTYE
PSSATPVTRSSQGSPSVVVAPSPKTKGQKAEDVPVRIALGASSCEICHEVFKSKNVRVLKCGHKYHKGCF
KQWLKGQSACPACQGRDLLTEESPSGRGWPSQNQELPSCSSR

Human TTC3 Protein sequence - var4 (public gi: 1632762) (SEQ ID NO: 311)
MDNFAEGDFTVADYALLEDCPHVDDCVFAAEFMSNDYVRVTQLYCDGVGVQYKDYIQSERNLEFDICSIW
CSKPISVLQDYCDAIKINIFWPLLFQHQNSSVISRLHPCVDANNSRASEINLKKLQHLELMEDIVDLAKK
VANDSFLIGGLLRIGCKIENKILAMEEALNWIKYAGDVTILTKLGSIDNCWPMLSIFFTEYKYHITKIVM
EDCNLLEELKTQSCMDCIEEGELMKMKGNEEFSKERFDIAIIYYTRAIEYRPENYLLYGNRALCFLRTGQ
FRNALGDGKRATILKNTWPKGHYRYCDALSMLGEYDWALQANIKAQKLCKNDPEGIKDLIQQHVKLQKQI
EDLQGRTANKDPIKAFYENRAYTPRSLSAPIFTTSLNFVEKERDFRKINHEMANGGNQNLKVADEALKVD
DCDCHPEFSPPSSQPPKHKGKQKSRNNESEKFSSSSPLTLPADLKNILEKQFSKSSRAAHQDFANIMKML
RSLIQDGYMALLEQRCRSAAQAFTELLNGLDPQKIKQLNLAMINYVLVVYGLAISLLGIGQPEELSEAEN
QFKRIIEHYPSEGLDCLAYCGIGKVYLKKNRFLEALNHFEKARTLIYRLPGVLTWPTSNVIIEESQPQKI
KMLLEKFVEECKFPPVPDAICCYQKCHGYSKIQIYITDPDFKGFIRISCCQYCKIEFHMNCWKKLKTTTF
NDKIDKDFLQGICLTPDCEGVISKIIIFSSGGEVKCEFEHKVIKEKVPPRPILKQKCSSLEKLRLKEDKK
LKRKIQKKEAKKLAQERMEEDLRESNPPKNEEQKETVDNVQRCQFLDDRILQCIKQYADKIKSGIQNTAM
LLKELLSWKVLSTEDYTTCFSSRNFLNEAVDYVIRHLIQENNRVKTRIFLHVLSELKEVEPKLAAWIQKL
NSFGLDATGTFFSRYGASLKLLDFSIMTFLWNEKYGHKLDSIEGKQLDYFSEPASLKEARCLIWLLEEHR
DKFPALHSALDEFFDIMDSRCTVLRKQDSGEAPFSSTKVKNKSKKKKPKDSKPMLVGSGTTSVTSNNEII
TSSEDHSNRNSDSAGPFAVPDHLRQDVEEFEALYDQHSNEYVVRNKKLWDMNPKQKCSTLYDYFSQFLEE
HGPLDMSNKMFSAEYEFFPEETRQILEKAGGLKPFLLGCPRFVVIDNCIALKKVASRLKKKRKKKNIKTK
VEEISKAGEYVRVKLQLNPAAREFKPDVKSKPVSDSSSAPAFENVKPKPVSANSPKPACEDVKAKPVSDN

SSRQVSEDGQPKGVSSNSPKPGSEDANYKRVSCNSPKPVLEDVKPTYWAQSHLVTGYCTYLPFQRFDITQ
TPPAYINVLPGLPQYTSIYTPLASLSPEYQLPRSVPVVPSFVANDRADKNAAAYFEGHHLNAENVAGHQI
ASETQILEGSLGISVKSHCSTGDAHTVLSESNRNDEHCGNSNNKCEVIPESTSAVTNIPHVQMVAIQVSW
NIIHQEVNTEPYNPFEERQGEISRIEKEHQVLQDQLQEVYENYEQIKLKGLEETRDLEEKLKRHLEENKI
SKTELDWFLQDLEREIKKWQQEKKEIQERLKSLKKKIKKVSNASEMYTQKNDGKEKEHELHLDQSLEISN
TLTNEKMKIEEYIKKGKEDYEESHQRAVAAEVSVLENWKESEVYKLQIMESQAEAFLKKLGLISRDPAAY
PDMESDIRSWELFLSNVTKEIEKAKSQFEEQIKAIKNGSRLSELSKVQISELSFPACNTVHPELLPESSG
HDGQGLVTSASDVTGNHAALHRDPSVFSAGDSPGEAPSALLPGPPPGQPEATQLTGPKRAGQAALSERSP
VADRKQPVPPGRAARSSQSPKKPFNSIIEHLSVVFPCYNSTELAGFIKKVRSKNKNSLSGLSIDEIVQRV
TEHILDEQKKKKPNPGKDKRTYEPSSATPVTRSSQGSPSVVVAPSPKTKGQKAEDVPVRIALGASSCEIC
HEVFKSKNVRVLKCGHKYHKGCFKQWLKGQSACPACQGRDLLTEESPSGRGWPSQNQELPSCSSR

Human TTC3 Protein sequence - var5 (public gi: 2969903) (SEQ ID NO: 312)
DLKKLQHLELMEDIVDLARKVANDSFLIGGLLRIGCKIENKILAMEEALNWIKYAGDVTILTKLGSIDNC
WPMLSIFFTEYKYHITKIVMEDCNLLEELKTQSCMDCIEEGGLMKMKGNEEFSKERFDIAIIYYTRAIEY
RPENYLLYGNRALCFPRTGQFRNALGDGKRATILKNTWPKGHYRYCDALSMLGEYDWALQANIKAQKLCK
NDPEGIKDLIQQHVKLQKQIEDLQGRTANKDPIKAFYENRAYTPRSLSAPIFTTSLNFVEKERDFRKINH
EMANGGNQNLKVADEALKVDDCDCHPEFSPPSSQPPKHKGKQKSRNNESEKFSSSSPLTLPADLKNILEK
QFSKSSRAAHQDFANIMKMLRSLIQDGYMALLEQRCRSAAQAFTELLNGLDPQKIKQLNLAMINYVLVVY
GLAISLLGIGQPEELSEAENQFKRIIEHYPSEGLDCLAYCGIGKVYLKKNRFLEALNHFEKARTLIYRLP
GVLTWP

Unigene Name: UBE2N **Unigene ID:** Hs.458359

Human UBE2N mRNA sequence - var1 (public gi: 37577134) (SEQ ID NO: 208)
CGCGCGCGCAGTCGCGCGCGGGTCGTGCCGTACCACCGTCGCGGGCAGGCTCGGCCACGAGCGCCAGAGC
CCCGCGCCTCCCCTCGCGGCCTGTCCCAAGTCCCTGCCCCGCAACAGAGCGTCACTTCCGCCATCCCCGG
CAGCGGTTGGGGCGGGGCGCACGGGGGAGGGGGCCAGGTCGGAGGGAAGCCCGCCCGTGCCCGAGCCCGC
GCCCGAGCAGGGACTACATTTCCCGAGGGGCCTCGGCGGCGGCTGCGGCGACGGGCGCGGCAACGTCCCC
CGGAAGTGGAGCCCGGGACTTCCACTCGTGCGTGAGGCGAGAGGAGCCGGAGACGAGACCAGAGGCCGAA
CTCGGGTTCTGACAAGATGGCCGGGCTGCCCCGCAGGATCATCAAGGAAACCCAGCGTTTGCTGGCAGAA
CCAGTTCCTGGCATCAAAGCCGAACCAGATGAGAGCAACGCCCGTTATTTTCATGTGGTCATTGCTGGCC
CTCAGGATTCCCCCTTTGAGGGAGGGACTTTTAAACTTGAACTATTCCTTCCAGAAGAATACCCAATGGC
AGCCCCTAAAGTACGTTTCATGACCAAAATTTATCATCCTAATGTAGACAAGTTGGGAAGAATATGTTTA
GATATTTTGAAAGATAAGTGGTCCCCAGCACTGCAGATCCGCACAGTTCTGCTATCGATCCAGGCCTTGT
TAAGTGCTCCCAATCCAGATGATCCATTAGCAAATGATGTAGCGGAGCAGTGGAAGACCAACGAAGCCCA
AGCCATAGAAACAGCTAGAGCATGGACTAGGCTATATGCCATGAATAATATTTAAATTGATACGATCATC
AAGTGTGCATCACTTCTCCTGTTCTGCCAAGACTTCCTCCTCTTTGTTTGCATTTAATGGACACAGTCTT
AGAAACATTACAGAATAAAAAAGCCCAGACATCTTCAGTCCTTTGGTGATTAAATGCACATTAGCAAATC
TATGTCTTGTCCTGATTCACTGTCATAAAGCATGAGCAGAGGCTAGAAGTATCATCTGGATTGTTGTGAA
ACGTTTAAAAGCAGTGGCCCCTCCCTGCTTTTATTCATTTCCCCCATCCTGGTTTAAGTATAAAGCACTG
TGAATGAAGGTAGTTGTCAGGTTAGCTGCAGGGGTGTGGGTGTTTTTATTTTATTTTATTTTATTTTATT
TTTGAGGGGGGAGGTAGTTTAATTTTATGGGCTCCTTTCCCCCTTTTTTGGTGATCTAATTGCATTGGTT
AAAAGCAGCTAACCAGGTCTTTAGAATATGCTCTAGCCAAGTCTAACTTTATTTAGACGCTGTAGATGGA
CAAGCTTGATTGTTGGAACCAAAATGGGAACATTAAACAAACATCACAGCCCTCACTAATAACATTGCTG
TCAAGTGTAGATTCCCCCCTTCAAAAAAAGCTTGTGACCATTTTGTATGGCTTGTCTGGAAACTTCTGTA
AATCTTATGTTTTAGTAAAATATTTTTTGTTATTCTACTTTGCCTTTGTACAGTTTATTTTACTGTGTTT
ATTTCATTTTCCCAATTTGACAATCGTATTTTAAAATTGAAACTGATGGAACATTCTTTCTTGGTCTTCA
CCATCTGACAAATTGAATGGCAAGAGGTGGATTTTGCCAGTTTCTTTTCACTGATGCAGATTTGTGTTAA
GATAGTACTGAATGGAGTATTTATAAACTGGCCCTGAGCATGCATAAAGCATCAGTATCTGACCTTTTTT
TAACCTTCTAGGAATTTGAAATAAATGTGTTTGTGTTGTCTGATTAGATGATCATTGGTGTCTTGCCACA
ATGTTTAAAAATTACTGTACAGGAAAGTCACAGCAAAGATAGCAGTTGTGACTGACATGTAGGACTTTCA
CAGTTGTGCCACATTTTTGCCTAAAATTTGGGTTATGACATTTTTCTTGGTTCTTATCTGAAAATTTCAT
CTGTAACCTTTCATGTGTGTTAAGAAACACTGATCTGATCATTTGGGATTTGCTGAGGCATTTGTGAGTC
TTCCTTATAAACCTGATGAGCAGATCTCAACTATCTAGCTTGTGTGTCATCAGAAAGGTTTATCCCTTTG
AGAGTATCAAGTCCTCAGTTAATGATTCTTGCTTTCATCCCTCCAGTATTTGCTGTGGGAGCTCGTTTTA
TTCTTTAATTTGGAATTCAGTAATTTTTCTTCTTTATTGACGAATTCCTCCCCTCACAAAACTGTTCTTT
CCCACCTCTCTCCATATCTAATTCCTGATTCTTGTTATTTTTAAGTCATAAATGTAGCCAGTCATAAATA
CATAAATGTTAACCTTCGGGTTGCAACCTTGTCTCTTGCAGTTTAAGGTAATGGATATTGTAGCCCATTT
GAATTTTCTTCACTCTTATTCTCGTAATTCTGGAGTTTCTTCAGATTGTGGTGTATTTTATTGTGCTCCT
ATGTAAGATGAAGAATTAACTATTAAAATTACATTTTCAACATACAAAAGCTTTTGATGACTGGTAACTG
GTATCCTTCCAAATAAATGCATTGCTTGGTAAAAAAAAAAAAAAAAAA

Human UBE2N protein sequence - var1 (public gi: 4507793) (SEQ ID NO: 313)
MAGLPRRIIKETQRLLAEPVPGIKAEPDESNARYFHVVIAGPQDSPFEGGTFKLELFLPEEYPMAAPKVR
FMTKIYHPNVDKLGRICLDILKDKWSPALQIRTVLLSIQALLSAPNPDDPLANDVAEQWKTNEAQAIETA
RAWTRLYAMNNI

Human UBE2N pray sequence - var1 (SEQ ID NO: 209)
GCCGCCATGGNGTACCCATACGACGTACCAGATTACGCTCATATGGCCATGGAGGCCAGTGAATTCCACC
CAAGCAGTGGTATCAACGCAGAGTGGCCGAACTCGGGTTCTGACAAGATGGCCGGGCTGCCCCGCAGGAT
CATCAAGGAAACCCAGCGTTTGCTGGCAGAACCAGTTCCTGGCATCAAAGCCGAACCAGATGAGAGCAAC
GCCCGTTATTTCATGTGGTCATTGCTGGCCCTCAGGATTCCCCCTTTGAGGGAGGGACTTTTAAACTTG
AACTATTCCTTCCAGAAGAATACCCAATGGCAGCCCCTAAAATAAGTGGTCCCCAGCACTGCAGATCCGC
ACAGTTCTGCTATCGATCCAGGCCTTGTTAAGTGCTCCCAATCCAGATGATCCATTAGCAAATGATGTAG
CGGAGCAGTGGAAGACCAACGAAGCCCAAGCCATAGAAACAGCTAGAGCATGGACTAGGCTATATGCCAT
GAATAATATTTAAATTGATACGATCATCAAGTGTGCATCACTTCTCCTGTTCTGCCAAGACTTCCTCCTC
TTTGTTTGCATTTAATGGACACAGTCTTAGAAACATTACAGAATAAAAANCCCCAGACATCTTCAGTCCT
TNGGTGATTAAATGCACATTANCAAATNTTGCCNTGTCCTGATNNCTGNCNTAANCNTGANCCNAGGCTN
AAATTTNATCTGGATNNTNGGAAACNTTNAAAACNNGGGCCCCNCCNGCTTTNTTNATTNCCCCANCCGG
NTNAANTTAAACCCNGGAATNANGGNNTTTNCNGNNACNNNGGGGGT

Human UBE2N pray sequence - var2 (SEQ ID NO: 210)
CGAGCGCCGCCTGGNNTACCCATACGACGTACCAGNATTACGCTCATATGGCCATGGNAGGCCAGTGAAT
TCCACCCAAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCCGGGGGAGAGGAGCCGGAGACGAGACCA
GAGGCCGAACTCGGGTTCTGACAAGATGGCCGGGCTGCCCCGCAGGATCATCAAGGAAACCCAGCGTTTG
CTGGCAGAACCAGTTCCTGGCATCAAAGCCGAACCAGATGAGAGCAACGCCCGTTATTTCATGTGGTCA
TTGCTGGCCCTCAGGATTCCCCCTTTGAGGGAGGGACTTTTAAACTTGAACTATTCCTTCCAGAAGAATA
CCCAATGGCAGCCCCTAAAGTACGTTTCATGACCAAAATTTATCATCCTAATGTAGACAAGTTGGGAAGA
ATATGTTTAGATATTTTGAAAGATAAGTGGTCCCCAGCACTGCAGATCCGCACAGTTCTGCTATCGATCC
AGGCCTTGTTAAGTGCTCCCAATCCAGATGATCCATTAGCAAATGATGTAGCGGAGCAGTGGAAGACCAA
CGAAGCCCAAGCCATAGAAACAGCTAGAGCATGGACTAGGCTATATGCCATGAATAATATTTAAATTGAT
ACGATCATCAAGTGTGCATCACTTCTCCTGTTCTGCCAAGACTTCNTCCTCTTTGTTTGCATTTAATGGA
CACAGTCTTANAAACNTTNNGAATAAAAANCCCANACNTTTNNNTCNTTNGTGATNAATGCCNTTANCAA
NNNTTNTNTTGNCNGNTNCCTGNNTAAACCTGNCCNAGNCTNAANTTNNNNNGGTTTTNNNAANNTTAAA
ANNNTGNCCCCCNNNTTTTTNTTNCCCCCNGNTNAANNAAANCNTNNANAANGNTTTTNNGTNNCCNN
GGGNGGGNTTTTTTTTTTTTTTTTNTN

Unigene Name: UNC84B **Unigene ID:** Hs.406612

Human UNC84B mRNA sequence - var1 (public gi: 31742497) (SEQ ID NO: 211)
CCGCCCGCCCGCCTTGTCCCGCGTCGCCGCTCTTCGCGTGCCCGCGCGCCCCGGGCCCGGCCGCTGTGTC
GCCCTGAGCGGAGCGCCGCCCGCGGGATCCCCACGCGGAAAGGGGGCGCGCCCCGGCGGCGGCCTGGCCT
CGGACGCCCCCGGCCGGGCTAGAAGCCGCCGCGGCAGCAGATTCTCTTCAGGGGAAGAGTCCACATCCCA
CCTCATCATGTCCCGAAGAAGCCAGCGCCTCACGCGCTACTCCCAGGGTGACGATGACGGCAGCAGCAGC
AGCGGAGGGAGCTCGGTGGCTGGGAGTCAGAGCACCCTGTTTAAAGACAGTCCTCTCAGGACCTTGAAGA
GGAAATCCAGCAACATGAAGCGCCTGTCCCCAGCGCCACAGCTGGGCCCGTCCTCTGATGCACACACCTC
CTACTACAGTGAGTCGCTGGTCCACGAGTCCTGGTTCCCACCCAGGAGCTCCCTGGAGGAACTGCATGGT
GACGCCAACTGGGGTGAGGACCTGCGGGTGCGGAGGAGGAGAGGCACGGGTGGCTCAGAGAGCAGCAGGG
CCAGCGGGCTTGTGGGGCGCAAGGCCACCGAGGACTTCCTGGGCTCTTCCTCGGGCTACTCCTCTGAGGA
CGACTACGTGGGCTACTCGGATGTGGACCAGCAGAGTTCCAGCTCGCGGCTCCGAAGCGCCGTCTCACGG
GCGGGCTCCTTACTCTGGATGGTGGCCACTTCGCCAGGCCGGCTCTTCAGACTTCTCTACTGGTGGGCTG
GCACCACCTGGTACCGCCTGACCACAGCTGCCTCCCTCCTTGACGTCTTCGTTTTAACCAGGCGCTTCTC
GTCCCTGAAGACGTTCCTCTGGTTCCTGCTGCCGCTGCTCTTGCTGACGTGCCTGACGTATGGTGCTTGG
TATTTCTACCCCTATGGGCTGCAGACATTCCACCCTGCTTTGGTTTCCTGGTGGGCAGCGAAGGACAGCA
GGAGGCCGGATGAGGGCTGGGAAGCCAGAGACTCATCGCCACATTTCCAGGCTGAGCAGCGTGTTATGTC
CCGGGTACACTCTCTGGAGCGGCGTCTGGAAGCTCTTGCTGCTGAATTTTCCTCCAACTGGCAGAAGGAG
GCCATGCGGCTGGAACGTCTGGAGCTGCGGCAAGGGGCTCCTGGCCAGGGAGGTGGTGGTGGCCTGAGCC
ACGAGGACACCCTGGCGCTGCTGGAGGGGCTAGTGAGCCGCCGTGAAGCTGCCCTGAAGGAGGATTTCCG
CAGGGAAACTGCTGCTCGCATCCAGGAAGAACTGTCTGCCCTGAGAGCAGAGCATCAGCAAGACTCAGAA
GACCTCTTCAAGAAGATCGTCCGGGCCTCCCAGGAGTCCGAGGCTCGCATCCAGCAGCTGAAGTCAGAGT

GGCAAAGCATGACCCAGGAGTCCTTCCAGGAGAGCTCTGTGAAGGAGCTGAGGCGGCTGGAGGACCAGCT
GGCCGGCCTGCAGCAGGAGCTGGCGGCTCTGGCACTGAAGCAGAGCTCGGTGGCGGAAGAAGTGGGCCTG
CTGCCCCAGCAGATCCAGGCCGTGCGGGACGACGTGGAATCTCAGTTCCCGGCCTGGATCAGTCAGTTCC
TTGCCCGAGGTGGAGGGGGCCGCGTGGGGCTCCTTCAGAGAGAGGAGATGCAAGCTCAGCTGCGAGAGCT
GGAGAGCAAGATCCTCACCCATGTGGCAGAGATGCAGGGCAAGTCGGCCAGGGAAGCCGCGGCCTCCCTG
AGCCTGACGCTGCAGAAAGAAGGTGTGATTGGAGTGACAGAGGAGCAGGTGCACCACATCGTGAAGCAGG
CCCTGCAGCGCTACAGTGAGGACCGCATCGGGCTGGCAGACTACGCCCTGGAGTCAGGAGGGGCCAGCGT
CATCAGCACCCGATGTTCTGAGACCTACGAGACCAAGACGGCCCTCCTCAGCCTCTTCGGCATCCCCCTG
TGGTACCACTCCCAGTCACCCCGAGTCATCCTCCAGCCAGATGTGCACCCAGGCAACTGCTGGGCCTTCC
AGGGGCCACAAGGCTTCGCCGTGGTCCGCCTCTCTGCCCGCATCCGCCCCACAGCCGTTACCTTAGAGCA
TGTGCCCAAGGCCTTGTCACCCAACAGCACTATCTCCAGTGCCCCCAAGGACTTCGCCATCTTTGGGTTT
GACGAAGACCTGCAGCAGGAGGGGACACTCCTTGGCAAGTTCACTTACGATCAGGACGGCGAGCCTATTC
AGACGTTTCACTTTCAGGCCCCTACGATGGCCACGTACCAGGTGGTGGAGCTGCGGATCCTGACTAACTG
GGGCCACCCCGAGTACACCTGCATCTACCGCTTCAGAGTGCATGGGGAGCCCGCCCACTAGCCCTGCTTA
CTGGTGCCTGCTGCCAGCCATCTGGGAGTGGGTGAACAGCACCCCGCCGCTTCCCCCACACGCTTGCTCG
GCGCTCTGACTTCTAGGAGCACAAGAGAGGAGCCTGTGGCCCCATGCAGATGAAAAGGACGGGCAGGGTC
TCCTGAGCAGCAGGTGGCTCGAGGCGGTTAGCAGGCTCCAGCAGCTCCCTTCTTCCTTCCCTCTGTGCCC
GTGGCGTCTGCTTCCCATCCTGGGAGTGTGTATATATGTAGCATATCATGGGGGACTGGGAAGTTGGGAG
AGGTAGGACCTGACTGGTCTTGGCTGGGGTCAGGGGCTGGTGCCTGGGAGCTGATGAAGCAGGTGCCAGG
GCTGTGGGAGGGGCAAGCTACGGCCTGGGCTAGGTGAGCTGCCTCTGCCCCTGGGCAAGGAAGCGAGGCC
CTCTGGGAGGAGGGTGCTTAGCTCCAGAGCAGGATGGGACTTCCCCAGGCAGGAAGCACTTGATGGAGAG
CTGCCCAGCTCTCCTACAAGGTTAGTGCCCTCCACCTAGGGAAGCATGAACCACAGGGTCCCTGAGGGCC
TTCGACAAAAGTGTGTATTTGTCCCGGGGAGGGTAGCAGTGGCCATGGGGCTTCTTGTGCCCTAAAGGG
GACTGGCTGCTGTGATCTTCTAAGGGGCCCAGGGCCAACCCTGTAGGCTTCCCCTCTGCTGGGGACGGTA
GTTGCTTTTCTCTCTCCTGATGCTAGGTTGGGGCCCACCCTGCTCCCTGTTCCTGCTAGGGCCTGCCAGT
GCCCCTGAGCTTGCTTTCCACATTCTCCCAGGGTATGGAGACCTAGACCTGTCTTTGGGGCCATTAGCAT
CTGGGGTTATAGCAAGAAGAGTGGGGAGCATGGAACTCCTGGGCTCTTGTGGGGACGTTCAGGGTATCGG
GGTGCGAGGTCTGTCTGCACCGGCCCCCACATCTAACCAGGCCCTGATGTAGGGGTCGTCCGCTCAGGCT
GCCCCCTTGGGCTCTTGCAGCTCTTGTTCAGGTAGTCGCCCTTCTGGTTTGTTCTCTGTGGGGCAGTTGG
TGGGGGCTGGGGAAGAGGCTGGCAGAAGTTACCCTGGATAGGGAAGGGGGAGGAGGGGACTTTTAGAGC
CAGCAGGCCCCACTGTATTATGTATATATTTTTCAAGGTCTGTTTTTCTAACTGAAAAGCTAAGGGCTTG
ATTCCTAGCCCCGTTCTGTGGGGCACTGGGTGATACTCAGTTTCTTGTTCCTGGCCGTGGAGAGGGGCCT
GGGGCACTGGTTCCGGCTGTGTCTGGTGGTCTGGCTGCGGGGAAGGGGCAAGAAGGCGGGCAGGCCTTCA
CTGCAGCACTGAGCCTCAAATCCGCTCTGGAGCATGAGGCTGGATGCAGTGGTGGTGAGGCCGCCCCGCT
CCATCCCGAGGCAGCCAGGGTTTGTTTTGCGCTCTCCTGTCACAAATGCTGCACTATTGGTTCTTAAGTT
TTTTATCTCCAGATCCTAATTTATGCCTATGCAAAAAATAAATGACGCCCAAGAGCTG

Human UNC84B protein sequence - var1 (public gi: 31742498) (SEQ ID NO: 314)
MSRRSQRLTRYSQGDDDGSSSSGGSSVAGSQSTLFKDSPLRTLKRKSSNMKRLSPAPQLGPSSDAHTSYY
SESLVHESWFPPRSSLEELHGDANWGEDLRVRRRRGTGGSESSRASGLVGRKATEDFLGSSSGYSSEDDY
VGYSDVDQQSSSSRLRSAVSRAGSLLWMVATSPGRLFRLLYWWAGTTWYRLTTAASLLDVFVLTRRFSSL
KTFLWFLLPLLLLTCLTYGAWYFYPYGLQTFHPALVSWWAAKDSRRPDEGWEARDSSPHFQAEQRVMSRV
HSLERRLEALAAEFSSNWQKEAMRLERLELRQGAPGQGGGGGLSHEDTLALLEGLVSRREAALKEDFRRE
TAARIQEELSALRAEHQQDSEDLFKKIVRASQESEARIQQLKSEWQSMTQESFQESSVKELRRLEDQLAG
LQQELAALALKQSSVAEEVGLLPQQIQAVRDDVESQFPAWISQFLARGGGRVGLLQREEMQAQLRELES
KILTHVAEMQGKSAREAAASLSLTLQKEGVIGVTEEQVHHIVKQALQRYSEDRIGLADYALESGGASVIS
TRCSETYETKTALLSLFGIPLWYHSQSPRVILQPDVHPGNCWAFQGPQGFAVVRLSARIRPTAVTLEHVP
KALSPNSTISSAPKDFAIFGFDEDLQQEGTLLGKFTYDQDGEPIQTFHFQAPTMATYQVVELRILTNWGH
PEYTCIYRFRVHGEPAH

Human UNC84B pray sequence - var1 (SEQ ID NO: 212)
GATTTGGNAATNCTACAGGGNATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATTCGATG
ATGAAGATACCCCACCAAACCCAAAAAAAGAGATCTTTAATACGACTCGACTATAGGGCGAGCGCCGCCA
TGGAGTACCCATACGACGTACCAGATTACGCTCATATGGCCATGGAGGCCAGTGAATTCCACCCAAGCAG
TGGTATCAACGCATAGTGGAAAAGCATGACCCAGGAGTCCTTCCAGGAGAGCTCTGTGAAGGAGCTGAGG
CGGCTGGAGGACCAGCTGGCCGGCCTGCAGCAGGAGCTGGCGGCTCTGGCACTGAAGCAGAGCTCGGTGG
CGGAAGAAGTGGGCCTGCTGCCCCAGCAGATCCAGGCCGTGCGGGACGACGTGGAATCTCAGTTCCCGGC
CTGGATCAGTCAGTTCCTTGCCCGAGGTGGAGGGGGCCGCGTGGGGCTCCTTCAGAGAGAGGAGATGCAA
GCTCAGCTGCGAGAGCTGGAGAGCAAGATCCTCACCCATGTGGCAGAGATGCAGGGCAAGTCGGCCAGGG
AAGCCGCGGCCTCCCTGAGCCTGACGCTNCANAAAGAAGGTGTGATTGGAGTGACAGAGGAGCAGGTGCA
CCACATCGTGAAGCAGGCCCTGCAGCGCTACAGTGAGGACCGCATCGGGCTGGCAGACTACGCCCTGGAG
TCAGGAGGGGCCAGCGTCATCAGCACCCGATGTTCTGAGACCTACNAGACCAAGACGGNCCTNCTCAGCC

TCTTNGGNATCCCCCTGGGGTACCACTCCCAGTCACCCCNAGTCATNCTCCANATGNGCACCCAGGCNAC
TGNTGGGCCTTNCAGGGGCCANNGGGNTTNNCCGGGGNCCGCNTTTTTCCCNA

Human UNC84B pray sequence - var2 (SEQ ID NO: 213)
CGCCGCCATGGTAGTACCCATACGACGTACCAGTATTACGCTCATATGGCTCATGGCAGGCCAGTGAATT
CCACCCAAGCAGTGGTATCAACGCAGAGTGGCCATTATGGCTCGGGGACGGCTGAGCCTATTCAGACGT
TTCACTTTCAGGCCCCTACGATGGCCACGTACCAGGTGGTGGAGCTGCGGATTCTGACTAACTGGGGCCA
CCCCGAGTACACCTGCATCTACCGCTTCAGAGTGCATGGGGAGCCCGCCCACTAGCCCTGCTTACTGGTG
CCCGCTGCCAGCCATCTGGGAGTGGGTGAACAGCACCCCGCCGCTTCCCCCACACGCTTGCTCGGCGCTC
TGACTTCTAGGAGCACAAGAGAGGAGCCTGTGGCCCCATGCAGATGAAAAGGACGGGCAGGGTCTCCTGA
GCANCAGGTGGCTCGAGGCGGTTAGCANGCTCCANCAGCTCCCTTCTTCCTTCCCTCTGTGCCCGTGGCG
TCTGCTTCCCATCCTGGGAGTGTGTNTATATNTANCATATCATGGGGGACTGG

Unigene Name: VCY2IP1 Unigene ID: Hs.66048 Clone ID: GD_181

Human VCY2IP1 mRNA sequence - var1 (public gi: 22002952) (SEQ ID NO: 214)
AAGATGGCGGCGGTGGCTGGATCTGGGGCTGCCGCGGCTCCGAGCTCACTGCTCCTCGTGGTGGGCAGCG
AGTTCGGGAGCCCGGGGCTCCTCACCTACGTCCTGGAGGAGCTCGAAAGAGGCATCCGGTCTTGGGATGT
CGATCCTGGCGTCTGCAACCTTGATGAACAGCTCAAGGTCTTTGTGTCCCGACACTCTGCCACCTTCTCC
AGCATTGTGAAAGGCCAGCGGAGCCTGCACCACCGTGGAGACAACCTGGAGACCCTGGTCCTCCTGAACC
CATCAGACAAGTCCCTGTATGATGAGCTCCGGAACCTTCTGTTGGACCCTGCCTCTCACAAGCTACTGGT
GTTGGCTGGGCTCTGCCTGGAGGAGACGGGGGAGCTGCTGCTACAGACAGGGGGCTTCTCGCCTCACCAC
TTCCTCCAGGTCCTGAAGGACAGAGAGATCCGGGACATCCTGGCCACCACGCCCCCACCTGTGCAGCCGC
CCATACTCACCATCACCTGCCCCACCTTCGGTGACTGGGCTCAGCCGGCACCCGCTGTGCCTGGCCTTCA
GGGGGCGCTCCGGCTCCAGCTGCGGCTGAACCCCCCGGCGCAGCTGCCCAACTCTGAGGGCCTGTGCGAA
TTCCTGGAGTACGTGGCTGAGTCTCTGGAGCCACCGTCCCCCTTCGAGCTGCTGGAGCCCCCGACCTCCG
GGGGCTTCCTCAGGCTGGGCCGGCCCTGCTGCTACATCTTCCCTGGAGGCCTCGGGGATGCCGCCTTCTT
CGCCGTCAATGGCTTCACTGTGCTGGTCAACGGTGGCTCAAACCCCAAGTCCAGTTTCTGGAAGCTGGTG
CGGCACCTGGACCGCGTGGATGCCGTGCTGGTGACCCACCCTGGCGCCGACAGCCTCCCCGGCCTCAACA
GCCTGCTGCGGCGCAAACTGGCGGAGCGCTCCGAGGTGGCTGCTGGTGGGGCTCCTGGGACGACAGGCT
GCGCAGGCTCATCTCCCCCAACCTGGGGGTCGTGTTCTTCAACGCCTGCGAGGCCGCGTCGCGGCTGGCG
CGCGGCGAGGATGAGGCGGAGCTGGCGCTGAGCCTCCTGGCGCAGCTGGGCATCACGCCTCTGCCACTCA
GCCGCGGCCCCGTGCCAGCCAAACCCACCGTGCTCTTCGAGAAGATGGGCGTGGGCCGGCTGGACATGTA
TGTGCTGCACCCGCCCTCCGCCGGCGCCGAGCGCACGCTGGCCTCTGTGTGCGCCCTGCTGGTGTGGCAC
CCCGCCGGCCCCGGCGAGAAGGTGGTGCGCGTGCTGTTCCCCGGTTGCACCCCGCCCGCCTGCCTCCTGG
ACGGCCTGGTCCGCCTGCAGCACTTGAGGTTCCTGCGAGAGCCCGTGGTGACGCCCCAGGACCTGGAGGG
GCCGGGGCGAGCCGAGAGCAAAGAGAGCGTGGGCTCCCGGGACAGCTCGAAGAGAGAGGGCCTCCTGGCC
ACCCACCCTAGACCTGGCCAGGAGCGCCCTGGGGTGGCCCGCAAGGAGCCAGCACGGGCTGAGGCCCCAC
GCAAGACTGAGAAAGAAGCCAAGACCCCCCGGGAGTTGAGGAAAGACCCCAAACCGAGTGTCTCCCGGAC
CCAGCCGCGGGAGGTGCGCCGGGCAGCCTCTTCTGTGCCCAACCTCAAGAAGACGAATGCCCAGGCGGCA
CCCAAGCCCCGCAAAGCGCCCAGCACGTCCCACTCTGGCTTCCCGCCGGTGGCAAATGGACCCCGCAGCC
CGCCCAGCCTCCGATGTGGAGAAGCCAGCCCCCCCAGTGCAGCCTGCGGCTCTCCGGCCTCCCAGCTGGT
GGCCACGCCCAGCCTGGAGCTGGGGCCGATCCCAGCCGGGGAGGAGAAGGCACTGGAGCTGCCTTTGGCC
GCCAGCTCAATCCCAAGGCCACGCACACCCTCCCCTGAGTCCCACCGGAGCCCCGCAGAGGGCAGCGAGC
GGCTGTCGCTGAGCCCACTGCGGGGCGGGGAGGCCGGGCCAGACGCCTCACCCACAGTGACCACACCCAC
GGTGACCACGCCCTCACTACCCGCAGAGGTGGGCTCCCCGCACTCGACCGAGGTGGACGAGTCCCTGTCG
GTGTCCTTTGAGCAGGTGCTGCCGCCATCCGCCCCCACCAGTGAGGCTGGGCTGAGCCTCCCGCTGCGTG
GCCCCCGGGCGCGGCGCTCGGCTTCCCCACACGATGTGGACCTGTGCCTGGTGTCACCCTGTGAATTTGA
GCATCGCAAGGCGGTGCCAATGGCACCGGCACCTGCGTCCCCCGGCAGCTCGAATGACAGCAGTGCCCGG
TCACAGGAACGGGCAGGTGGGCTGGGGGCCGAGGAGACGCCACCCACATCGGTCAGCGAGTCCCTGCCCA
CCCTGTCTGACTCGGATCCCGTGCCCCTGGCCCCCGGTGCGGCAGACTCAGACGAAGACACAGAGGGCTT
TGGAGTCCCTCGCCACGACCCTTTGCCTGACCCCCTCAAGGTCCCCCCACCACTGCCTGACCCATCCAGC
ATCTGCATGGTGGACCCCGAGATGCTGCCCCCCAAGACAGCACGGCAAACGGAGAACGTCAGCCGCACCC
GGAAGCCCCTGGCCCGCCCCAACTCACGCGCTGCCGCCCCCAAAGCCACTCCAGTGGCTGCTGCCAAAAC
CAAGGGGCTTGCTGGTGGGGACCGTGCCAGCCGACCACTCAGTGCCCGGAGTGAGCCCAGTGAGAAGGGA
GGCCGGGCACCCCTGTCCAGAAAGTCCTCAACCCCCAAGACTGCCACTCGAGGCCCGTCGGGGTCAGCCA
GCAGCCGGCCCGGGGTGTCAGCCACCCCACCCAAGTCCCCGGTCTACCTGGACCTGGCCTACCTGCCCAG
CGGGAGCAGCGCCCACCTGGTGGATGAGGAGTTCTTCCAGCGCGTGCGCGCGCTCTGCTACGTCATCAGT
GGCCAGGACCAGCGCAAGGAGGAAGGCATGCGGGCCGTCCTGGACGCGCTACTGGCCAGCAAGCAGCATT
GGGACCGTGACCTGCAGGTGACCCTGATCCCCACTTTCGACTCGGTGGCCATGCATACGTGGTACGCAGA
GACGCACGCCCGGCACCAGGCGCTGGGCATCACGGTGTTGGCAGCAACAGCATGGTGTCCATGCAGGAT
GACGCCTTCCCGGCCTGCAAGGTGGAGTTCTAGCCCCATCGCCGACACGCCCCCCACTCAGCCCAGCCCG

CCTGTCCCTAGATTCAGCCACATCAGAAATAAACTGTGACTAC

Human VCY2IP1 mRNA sequence - var2 (public gi: 21739762) (SEQ ID NO: 215)
CCGAAGATGGCGGCGGTGGCTGGATCTGGGGCTGCCGCGGCTCCGAGCTCACTGCTCCTCGTGGTGGGCA
GCGAGTTCGGGAGCCCGGGGCTCCTCACCTACGTCCTGGAGGAGCTCGAAAGAGGCATCCGGTCTTGGGA
TGTCGATCCTGGCGTCTGCAACCTTGATGAACAGCTCAAGGTCTTTGTGTCCCGACACTCTGCCACCTTC
TCCAGCATTGTGAAAGGCCAGCGGAGCCTGCACCACCGTGGAGACAACCTGGAGACCCTGGTCCTCCTGA
ACCCATCAGACAAGTCCCTGTATGATGAGCTCCGGAACCTTCTGTTGGACCCTGCCTCTCACAAGCTACT
GGTGTTGGCTGGGCCCTGCCTGGAGGAGACGGGGGAGCTGCTGCTACAGACAGGGGGCTTCTCGCCTCAC
CACTTCCTCCAGGTCCTGAAGGACAGAGAGATCCGGGACATCCTGGCCACCACGCCCCCACCTGTGCAGC
CGCCCATACTCACCATCACCTGCCCCACCTTCGGTGACTGGGCTCAGCTGGCACCCGCTGTGCCTGGCCT
TCAGGGGGCGCTCCGGCTCCAGCTGCGGCTGAACCCCCCGGCGCAGCTGCCCAACTCTGAGGGCCTGTGC
GAATTCCTGGAGTACGTGGCTGAGTCTCTGGAGCCACCGTCCCCCTTCGAGCTGCTGGAGCCCCCGACCT
CCGGGGGCTTCCTCAGGCTGGGCCGGCCCTGCTGCTACATCTTCCCTGGAGGCCTCGGGGATGCCGCCTT
CTTCGCCGTCAATGGCTTCACTGTGCTGGTCAACGGTGGCTCAAACCCCAAGTCCAGTTTCTGGAAGCTG
GTGCGGCACCTGGACCGCGTGGATGCCGTGCTGGTGACCCACCCTGGCGCCGACAGCCTCCCTGGCCTCA
ACAGCCTGCTGCGGCGCAAACTGGCGGAGCGCTCCGAGGTGGCTGCTGGTGGGGGCTCCTGGGACGACAG
GCTGCGCAGGCTCATCTCCCCCAACCTGGGGGTCGTGTTCTTCAACGCCTGCGAGGCCGCGTCGCGGCTG
GCGCGCGGCGAGGATGAGGCGGAGCTGGCGCTGAGCCTCCTGGCGCAGCTGGGCATCACGCCTCTGCCAC
TCAGCCGCGGCCCCGTGCCAGCCAAACCCACCGTGCTCTTCGAGAAGATGGGCGTGGGCCGGCTGGACAT
GTATGTGCTGCACCCGCCCTCCGCCGGCGCCGAGCGCACGCTGGCCTCTGTGTGCGCCCTGCTGGTGTGG
CACCCCGCCGGCCCCGGCGAGAAGGTGGTGCGCGTGCTGTTCCCCGGTTGCACCCCGCCCGCCTACCTCC
TGGACGGCCTGGTCCGCCTGCAGCACTTGAGGTTCCTGCGAGAGCCCGTGGTGACGCCCCAGGACCTGGA
GGGGCCGGGGCGAGCCGAGAGCAAAGAGAGCGTGGGCTCCCGGGACAGCTCGAAGAGAGAGGGCCTCCTG
GCCACCCACCCTAGACCTGGCCAGGAGCGCCCTGGGGTGGCCCGCAAGGAGCCAGCACGGGCTGAGGCCC
CACGCAAGACTGAGAAAGAAGCCAAGACCCCCCGGGAGTTGAAGAAAGACCCCAAACCGAGTGTCTCCCG
GACCCAGCCGCGGGAGGTGCGCCGGGCAGCCTCTTCTGTGCCCAACCTCAAGAAGACGAATGCCCAGGCG
GCACCCAAGCCCCGCAAAGCGCCCAGCACGTCCCACTCTGGCTTCCCGCCGGTGGCAAATGGACCCCGCA
GCCCGCCCAGCCTCCGATGTGGAGAAGCCAGCCCCCCCAGTGCAGCCTGCGGCTCTCCGGCCTCCCAGCT
GGTGGCCACGCCCAGCCTGGAGCTGGGGCCGATCCCAGCCGGGGAGGAGAAGGCACTGGAGCTGCCTTTG
GCCGCCAGCTCAATCCCAAGGCCACGCACACCCTCCCCTGAGTCCCACCGGAGCCCCGCAGAGGGCAGCG
AGCGGCTGTCGCTGAGCCCACTGCGGGGCGGGGAGGCCGGGCCAGACGCCTCACCCACAGTGACCACACC
CACGGTGACCACGCCCTCACTACCCGCAGAGGTGGGCTCCCCGCACTCGACCGAGGTGGACGAGTCCCTG
TCGGTGTCCTTTGAGCAGGTGCTGCCGCCATCCGCCCCCACCAGTGAGGCTGGGCTGAGCCTCCCGCTGC
GTGGCCCCCGGGCGCGGCGCTCGGCTTCCCCACACGATGTGGACCTGTGCCTGGTGTCACCCTGTGAATT
TGAGCATCGCAAGGCGGTGCCAATGGCACCGGCACCTGCGTCCCCCGGCAGCTCGAATGACAGCAGTGCC
CGGTCACAGGAACGGGCAGGTGGGCTGGGGGCCGAGGAGACGCCACCCACATCGGTCAGCGAGTCCCTGC
CCACCCTGTCTGACTCGGATCCCGTGCCCCTGGCCCCCGGTGCGGCAGACTCAGACGAAGACACAGAGGG
CTTTGGAGTCCCTCGCCACGACCCTTTGCCTGACCCCCTCAAGGTCCCCCCACCACTGCCTGACCCATCC
AGCATCTGCATGGTGGACCCCGAGATGCTGCCCCCCAAGACAGCACGGCAAACGGAGAACGTCAGCCGCA
CCCGGAAGCCCCTGGCCCGCCCCAACTCACGCGCTGCCGCCCCCAAAGCCACTCCAGTGGCTGCTGCCAA
AACCAAGGGGCTTGCTGGTGGGGACCGTGCCAGCCGACCACTCAGTGCCCGGAGTGAGCCCAGTGAGAAG
GGAGGCCGGGCACCCCTGTCCAGAAAGTCCTCAACCCCCAAGACTGCCACTCGAGGCCCGTCGGGGTCAG
CCAGCAGCCGGCCCGGGGTGTCAGCCACCCCACCCAAGTCCCCGGTCTACCTGGACCTGGCCTACCTGCC
CAGCGGGAGCAGCGCCCACCTGGTGGATGAGGAGTTCTTCCAGCGCGTGCGCGCGCTCTGCTACGTCATC
AGTGGCCAGGACCAGCGCAAGGAGGAAGGCATGCGGGCCGTCCTGGACGCGCTACTGGCCAGCAAGCAGC
ATTGGGACCGTGACCTGCAGGTGACCCTGATCCCCACTTTCGACTCGGTGGCCATGCATACGTGGTACGC
AGAGACGCACGCCCGGCACCAGGCGCTGGGCATCACGGTGTTGGGCAGCAACAGCATGGTGTCCATGCAG
GATGACGCCTTCCCGGCCTGCAAGGTGGAGTTCTAGCCCCATCGCCGACACGCCCCCCACTCAGCCCAGC
CCGCCTGTCCCTAGATTCAGCCACATCAGAAATAAACTGTGACTACACTTGGTAAAAAAAAAAAAAAAAA
AA

Human VCY2IP1 mRNA sequence - var3 (public gi: 21104445) (SEQ ID NO: 216)
CCGAGGTGGCTGCTGGTGGGGGCTCCTGGGACGACAGGCTGCGCAGGCTCATCTCCCCCAACCTGGGGGT
CGTGTTCTTCAACGCCTGCGAGGCCGCGTCGCGGCTGGCGCGCGGCGAGGATGAGGCGGAGCTGGCGCTG
AGCCTCCTGGCGCAGCTGGGCATCACGCCTCTGCCACTCAGCCGCGGCCCCGTGCCAGCCAAACCCACCG
TGCTCTTCGAGAAGATGGGCGTGGGCCGGCTGGACATGTATGTGCTGCACCCGCCCTCCGCCGGCGCCGA
GCGCACGCTGGCCTCTGTGTGCGCCCTGCTGGTGTGGCACCCCGCCGGCCCCGGCGAGAAGGTGGTGCGC
GTGCTGTTCCCCGGTTGCACCCCGCCCGCCTGCCTCCTGGACGGCCTGGTCCGCCTGCAGCACTTGAGGT
TCCTGCGAGAGCCCGTGGTGACGCCCCAGGACCTGGAGGGGCCGGGGCGAGCCGAGAGCAAAGAGAGCGT
GGGCTCCCGGGACAGCTCGAAGAGAGAGGGCCTCCTGGCCACCCACCCTAGACCTGGCCAGGAGCGCCCT
GGGGTGGCCCGCAAGGAGCCAGCACGGGCTGAGGCCCCACGCAAGACTGAGAAAGAAGCCAAGGCCCCCC

GGGAGTTGAAGAAAGACCCCAAACCGAGTGTCTCCCGGACCCAGCCGCGGGAGGTGCGCCGGGCAGCCTC
TTCTGTGCCCAACCTCAAGAAGACGAATGCCCAGGCGGCACCCAAGCCCCGCAAAGCGCCCAGCACGTCC
CACTCTGGCTTCCCGCCGGTGGCAAATGGACCCCGCAGCCCGCCCAGCCTCCGATGTGGAGAAGCCAGCC
CCCCCAGTGCAGCCTGCGGCTCTCCGGCCTCCCAGCTGGTGGCCACGCCCAGCCTGGAGCTGGGGCCGAT
CCCAGCCGGGGAGGAGAAGGCACTGGAGCTGCCTTTGGCCGCCAGCTCAATCCCAAGGCCACGCACACCC
TCCCCTGAGTCCCACCGGAGCCCCGCAGAGGGCAGCGAGCGGCTGTCGCTGAGCCCACTGCGGGGCGGGG
AGGCCGGGCCAGACGCCTCACCCACAGTGACCACACCCACGGTGACCACGCCCTCACTACCCGCAGAGGT
GGGCTCCCCGCACTCGACCGAGGTGGACGAGTCCCTGTCGGTGTCCTTTGAGCAGGTGCTGCCGCCATCC
GCCCCCACCAGTGAGGCTGGGCTGAGCCTCCCGCTGCGTGGCCCCCGGGCGCGGCGCTCGGCTTCCCCAC
ACGATGTGGACCTGTGCCTGGTGTCACCCTGTGAATTTGAGCATCGCAAGGCGGTGCCAATGGCACCGGC
ACCTGCGTCCCCCGGCAGCTCGAATGACAGCAGTGCCCGGTCACAGGAACGGGCAGGTGGGCTGGGGGCC
GAGGAGACGCCACCCACATCGGTCAGCGAGTCCCTGCCCACCCTGTCTGACTCGGATCCCGTGCCCCTGG
CCCCCGGTGCGGCAGACTCAGACGAAGACACAGAGGGCTTTGGAGTCCCTCGCCACGACCCTTTGCCTGA
CCCCCTCAAGGTCCCCCCACCACTGCCTGACCCATCCAGCATCTGCATGGTGGACCCCGAGATGCTGCCC
CCCAAGACAGCACGGCAAACGGAGAACGTCAGCCGCACCCGGAAGCCCCTGGCCCGCCCCAACTCACGCG
CTGCCGCCCCCAAAGCCACTCCAGTGGCTGCTGCCAAAACCAAGGGGCTTGCTGGTGGGGACCGTGCCAG
CCGACCACTCAGTGCCCGGAGTGAGCCCAGTGAGAAGGGAGGCCGGGCACCCCTGTCCAGAAAGTCCTCA
ACCCCCAAGACTGCCACTCGAGGCCCGTCGGGGTCAGCCAGCAGCCGGCCCGGGGTGTCAGCCACCCCAC
CCAAGTCCCCGGTCTACCTGGACCTGGCCTACCTGCCCAGCGGGAGCAGCGCCCACCTGGTGGATGAGGA
GTTCTTCCAGCGCGTGCGCGCGCTCTGCTACGTCATCAGTGGCCAGGACCAGCGCAAGGAGGAAGGCATG
CGGGCCGTCCTGGACGCGCTACTGGCCAGCAAGCAGCATTGGGACCGTGACCTGCAGGTGACCCTGATCC
CCACTTTCGACTCGGTGGCCATGCATACGTGGTACGCAGAGACGCACGCCCGGCACCAGGCGCTGGGCAT
CACGGTGTTGGGCAGCAACAGCATGGTGTCCATGCAGGATGACGCCTTCCCGGCCTGCAAGGTGGAGTTC
TAGCCCCATCGCCGACACGCCCCCCACTCAGCCCAGCCCGCCTGTCCCTAGATTCAGCCACATCAGAAAT
AAACTGTGACTTCCAAAAAAAAAAA

Human VCY2IP1 mRNA sequence - var4 (public gi: 14250679) (SEQ ID NO: 217)

GGCACGAGGCCGCCTTCTTCGCCGTCAATGGCTTCACTGTGCTGGTCAACGGTGGCTCAAACCCCAAGTC
CAGTTTCTGGAAGCTGGTGCGGCACCTGGACCGCGTGGATGCCGTGCTGGTGACCCACCCTGGCGCCGAC
AGCCTCCCCGGCCTCAACAGCCTGCTGCGGCGCAAACTGGCGGAGCGCTCCGAGGTGGCTGCTGGTGGGG
GCTCCTGGGACGACAGGCTGCGCAGGCTCATCTCCCCCAACCTGGGGGTCGTGTTCTTCAACGCCTGCGA
GGCCGCGTCGCGGCTGGCGCGCGGCGAGGATGAGGCGGAGCTGGCGCTGAGCCTCCTGGCGCAGCTGGGC
ATCACGCCTCTGCCACTCAGCCGCGGCCCCGTGCCAGCCAAACCCACCGTGCTCTTCGAGAAGATGGGCG
TGGGCCGGCTGGACATGTATGTGCTGCACCCGCCCTCCGCCGGCGCCGAGCGCACGCTGGCCTCTGTGTG
CGCCCTGCTGGTGTGGCACCCCGCCGGCCCCGGCGAGAAGGTGGTGCGCGTGCTGTTCCCCGGTTGCACC
CCGCCCGCCTGCCTCCTGGACGGCCTGGTCCGCCTGCAGCACTTGAGGTTCCTGCGAGAGCCCGTGGTGA
CGCCCCAGGACCTGGAGGGGCCGGGGCGAGCCGAGAGCAAAGAGAGCGTGGGCTCCCGGGACAGCTCGAA
GAGAGAGGGCCTCCTGGCCACCCACCCTAGACCTGGCCAGGAGCGCCCTGGGGTGGCCCGCAAGGAGCCA
GCACGGGCTGAGGCCCCACGCAAGACTGAGAAAGAAGCCAAGACCCCCCGGGAGTTGAAGAAAGACCCCA
AACCGAGTGTCTCCCGGACCCAGCCGCGGGAGGTGCGCCGGGCAGCCTCTTCTGTGCCCAACCTCAAGAA
GACGAATGCCCAGGCGGCACCCAAGCCCCGCAAAGCGCCCAGCACGTCCCACTCTGGCTTCCCGCCGGTG
GCAAATGGACCCCGCAGCCCGCCCAGCCTCCGATGTGGAGAAGCCAGCCCCCCCAGTGCAGCCTGCGGCT
CTCCGGCCTCCCAGCTGGTGGCCACGCCCAGCCTGGAGCTGGGGCCGATCCCAGCCGGGGAGGAGAAGGC
ACTGGAGCTGCCTTTGGCCGCCAGCTCAATCCCAAGGCCACGCACACCCTCCCCTGAGTCCCACCGGAGC
CCCGCAGAGGGCAGCGAGCGGCTGTCGCTGAGCCCACTGCGGGGCGGGGAGGCCGGGCCAGACGCCTCAC
CCACAGTGACCACACCCACGGTGACCACGCCCTCACTACCCGCAGAGGTGGGCTCCCCGCACTCGACCGA
GGTGGACGAGTCCCTGTCGGTGTCCTTTGAGCAGGTGCTGCCGCCATCCGCCCCCACCAGTGAGGCTGGG
CTGAGCCTCCCGCTGCGTGGCCCCCGGGCGCGGCGCTCGGCTTCCCCACACGATGTGGACCTGTGCCTGG
TGTCACCCTGTGAATTTGAGCATCGCAAGGCGGTGCCAATGGCACCGGCACCTGCGTCCCCCGGCAGCTC
GAATGACAGCAGTGCCCGGTCACAGGAACGGGCAGGTGGGCTGGGGGCCGAGGAGACGCCACCCACATCG
GTCAGCGAGTCCCTGCCCACCCTGTCTGACTCGGATCCCGTGCCCCTGGCCCCCGGTGCGGCAGACTCAG
ACGAAGACACAGAGGGCTTTGGAGTCCCTCGCCACGACCCTTTGCCTGACCCCCTCAAGGTCCCCCCACC
ACTGCCTGACCCATCCAGCATCTGCATGGTGGACCCCGAGATGCTGCCCCCCAAGACAGCACGGCAAACG
GAGAACGTCAGCCGCACCCGGAAGCCCCTGGCCCGCCCCAACTCACGCGCTGCCGCCCCCAAAGCCACTC
CAGTGGCTGCTGCCAAAACCAAGGGGCTTGCTGGTGGGGACCGTGCCAGCCGACCACTCAGTGCCCGGAG
TGAGCCCAGTGAGAAGGGAGGCCGGGCACCCCTGTCCAGAAAGTCCTCAACCCCCAAGACTGCCACTCGA
GGCCCGTCGGGGTCAGCCAGCAGCCGGCCCGGGGTGTCAGCCACCCCACCCAAGTCCCCGGTCTACCTGG
ACCTGGCCTACCTGCCCAGCGGGAGCAGCGCCCACCTGGTGGATGAGGAGTTCTTCCAGCGCGTGCGCGC
GCTCTGCTACGTCATCAGTGGCCAGGACCAGCGCAAGGAGGAAGGCATGCGGGCCGTCCTGGACGCGCTA
CTGGCCAGCAAGCAGCATTGGGACCGTGACCTGCAGGTGACCCTGATCCCCACTTTCGACTCGGTGGCCA
TGCATACGTGGTACGCAGAGACGCACGCCCGGCACCAGGCGCTGGGCATCACGGTGTTGGGCAGCAACAG
CATGGTGTCCATGCAGGATGACGCCTTCCCGGCTTGCAAGGTGGAGTTCTAGCCCCATCGCCGACACGCC
CCCCACTCAGCCCAGCCCGCCTGTCCCTAGATTCAGCCACATCAGAAATAAACTGTGACTACACTTGAAA

AAAAAAAAAAAAAAAAAAAAAAAA

Human VCY2IP1 mRNA sequence - var5 (public gi: 13938254) (SEQ ID NO: 218)

GACACCGACAGGGACTCGTCCACCTCGGTGTCCTTTGAGCAGGTGCTGCCGCCATCCGCCCCCACCAGTG
AGGCTGGGCTGAGCCTCCCGCTGCGTGGCCCCCGGGCGCGGCGCTCGGCTTCCCCACACGATGTGGACCT
GTGCCTGGTGTCACCCTGTGAATTTGAGCATCGCAAGGCGGTGCCAATGGCACCGGCACCTGCGTCCCCC
GGCAGCTCGAATGACAGCAGTGCCCGGTCACAGGAACGGGCAGGTGGGCTGGGGGCCGAGGAGACGCCAC
CCACATCGGTCAGCGAGTCCCTGCCCACCCTGTCTGACTCGGATCCCGTGCCCCTGGCCCCCGGTGCGGC
AGACTCAGACGAAGACACAGAGGGCTTTGGAGTCCCTCGCCACGACCCTTTGCCTGACCCCCTCAAGGTC
CCCCCACCACTGCCTGACCCATCCAGCATCTGCATGGTGGACCCCGAGATGCTGCCCCCCAAGACAGCAC
GGCAAACGGAGAACGTCAGCCGCACCCGGAAGCCCCTGGCCCGCCCCAACTCACGCGCTGCCGCCCCCAA
AGCCACTCCAGTGGCTGCTGCCAAAACCAAGGGGCTTGCTGGTGGGGACCGTGCCAGCCGACCACTCAGT
GCCCGGAGTGAGCCCAGTGAGAAGGGAGGCCGGGCACCCCTGTCCAGAAAGTCCTCAACCCCCAAGACTG
CCACTCGAGGCCCGTCGGGGTCAGCCAGCAGCCGGCCCGGGGTGTCAGCCACCCCACCCAAGTCCCCGGT
CTACCTGGACCTGGCCTACCTGCCCAGCGGGAGCAGCGCCCACCTGGTGGATGAGGAGTTCTTCCAGCGC
GTGCGCGCGCTCTGCTACGTCATCAGTGGCCAGGACCAGCGCAAGGAGGAAGGCATGCGGGCCGTCCTGG
ACGCGCTACTGGCCAGCAAGCAGCATTGGGACCGTGACCTGCAGGTGACCCTGATCCCCACTTTCGACTC
GGTGGCCATGCATACGTGGTACGCAGAGACGCACGCCCGGCACCAGGCGCTGGGCATCACGGTGTTGGGC
AGCAACAGCATGGTGTCCATGCAGGATGACGCCTTCCCGGCCTGCAAGGTGGAGTTCTAGCCCCATCGCC
GACACGCCCCCCACTCAGCCCAGCCCGCCTGTCCCTAGATTCAGCCACATCAGAAATAAACTGTGACTAC
ACTTAAAAAAAAAAAAAAAAAAAAAA

Human VCY2IP1 mRNA sequence - var6 (public gi: 14042428) (SEQ ID NO: 219)

AAGATGGCGGCGGTGGCTGGATCTGGGGCTGCCGCGGCTCCGAGCTCACTGCTCCTCGTGGTGGGCAGCG
AGTTCGGGAGCCCGGGGCTCCTCACCTACGTCCTGGAGGAGCTCGAAAGAGGCATCCGGTCTTGGGATGT
CGATCCTGGCGTCTGCAACCTTGATGAACAGCTCAAGGTCTTTGTGTCCCGACACTCTGCCACCTTCTCC
AGCATTGTGAAAGGCCAGCGGAGCCTGCACCACCGTGGAGACAACCTGGAGACCCTGGTCCTCCTGAACC
CATCAGACAAGTCCCTGTATGATGAGCTCCGGAACCTTCTGTTGGACCCTGCCTCTCACAAGCTACTGGT
GTTGGCTGGGCTCTGCCTGGAGGAGACGGGGGAGCTGCTGCTACAGACAGGGGGCTTCTCGCCTCACCAC
TTCCTCCAGGTCCTGAAGGACAGAGAGATCCGGGACATCCTGGCCACCACGCCCCCACCTGTGCAGCCGC
CCATACTCACCATCACCTGCCCCACCTTCGGTGACTGGGCTCAGCCGGCACCCGCTGTGCCTGGCCTTCA
GGGGGCGCTCCGGCTCCAGCTGCGGCTGAACCCCCCGGCGCAGCTGCCCAACTCTGAGGGCCTGTGCGAA
TTCCTGGAGTACGTGGCTGAGTCTCTGGAGCCACCGTCCCCCTTCGAGCTGCTGGAGCCCCCGACCTCCG
GGGGCTTCCTCAGGCTGGGCCGGCCCTGCTGCTACATCTTCCCTGGAGGCCTCGGGGATGCCGCCTTCTT
CGCCGTCAATGGCTTCACTGTGCTGGTCAACGGTGGCTCAAACCCCAAGTCCAGTTTCTGGAAGCTGGTG
CGGCACCTGGACCGCGTGGATGCCGTGCTGGTGACCCACCCTGGCGCCGACAGCCTCCCCGGCCTCAACA
GCCTGCTGCGGCGCAAACTGGCGGAGCGCTCCGAGGTGGCTGCTGGTGGGGGCTCCTGGGACGACAGGCT
GCGCAGGCTCATCTCCCCCAACCTGGGGGTCGTGTTCTTCAACGCCTGCGAGGCCGCGTCGCGGCTGGCG
CGCGGCGAGGATGAGGCGGAGCTGGCGCTGAGCCTCCTGGCGCAGCTGGGCATCACGCCTCTGCCACTCA
GCCGCGGCCCCGTGCCAGCCAAACCCACCGTGCTCTTCGAGAAGATGGGCGTGGGCCGGCTGGACATGTA
TGTGCTGCACCCGCCCTCCGCCGGCGCCGAGCGCACGCTGGCCTCTGTGTGCGCCCTGCTGGTGTGGCAC
CCCGCCGGCCCCGGCGAGAAGGTGGTGCGCGTGCTGTTCCCCGGTTGCACCCCGCCCGCCTGCCTCCTGG
ACGGCCTGGTCCGCCTGCAGCACTTGAGGTTCCTGCGAGAGCCCGTGGTGACGCCCCAGGACCTGGAGGG
GCCGGGGCGAGCCGAGAGCAAAGAGAGCGTGGGCTCCCGGGACAGCTCGAAGAGAGAGGGCCTCCTGGCC
ACCCACCCTAGACCTGGCCAGGAGCGCCCTGGGGTGGCCCGCAAGGAGCCAGCACGGGCTGAGGCCCCAC
GCAAGACTGAGAAAGAAGCCAAGACCCCCCGGGAGTTGAGGAAAGACCCCAAACCGAGTGTCTCCCGGAC
CCAGCCGCGGGAGGTGCGCCGGGCAGCCTCTTCTGTGCCCAACCTCAAGAAGACGAATGCCCAGGCGGCA
CCCAAGCCCCGCAAAGCGCCCAGCACGTCCCACTCTGGCTTCCCGCCGGTGGCAAATGGACCCCGCAGCC
CGCCCAGCCTCCGATGTGGAGAAGCCAGCCCCCCCAGTGCAGCCTGCGGCTCTCCGGCCTCCCAGCTGGT
GGCCACGCCCAGCCTGGAGCTGGGGCCGATCCCAGCCGGGGAGGAGAAGGCACTGGAGCTGCCTTTGGCC
GCCAGCTCAATCCCAAGGCCACGCACACCCTCCCCTGAGTCCCACCGGAGCCCCGCAGAGGGCAGCGAGC
GGCTGTCGCTGAGCCCACTGCGGGGCGGGGAGGCCGGGCCAGACGCCTCACCCACAGTGACCACACCCAC
GGTGACCACGCCCTCACTACCCGCAGAGGTGGGCTCCCCGCACTCGACCGAGGTGGACGAGTCCCTGTCG
GTGTCCTTTGAGCAGGTGCTGCCGCCATCCGCCCCCACCAGTGAGGCTGGGCTGAGCCTCCCGCTGCGTG
GCCCCCGGGCGCGGCGCTCGGCTTCCCCACACGATGTGGACCTGTGCCTGGTGTCACCCTGTGAATTTGA
GCATCGCAAGGCGGTGCCAATGGCACCGGCACCTGCGTCCCCCGGCAGCTCGAATGACAGCAGTGCCCGG
TCACAGGAACGGGCAGGTGGGCTGGGGGCCGAGGAGACGCCACCCACATCGGTCAGCGAGTCCCTGCCCA
CCCTGTCTGACTCGGATCCCGTGCCCCTGGCCCCCGGTGCGGCAGACTCAGACGAAGACACAGAGGGCTT
TGGAGTCCCTCGCCACGACCCTTTGCCTGACCCCCTCAAGGTCCCCCCACCACTGCCTGACCCATCCAGC
ATCTGCATGGTGGACCCCGAGATGCTGCCCCCCAAGACAGCACGGCAAACGGAGAACGTCAGCCGCACC
CGGAAGCCCCTGGCCCGCCCCAACTCACGCGCTGCCGCCCCCAAAGCCACTCCAGTGGCTGCTGCCAAAA
CCAAGGGGCTTGCTGGTGGGGACCGTGCCAGCCGACCACTCAGTGCCCGGAGTGAGCCCAGTGAGAAGGG

AGGCCGGGCACCCCTGTCCAGAAAGTCCTCAACCCCCAAGACTGCCACTCGAGGCCCGTCGGGGTCAGCC
AGCAGCCGGCCCGGGGTGTCAGCCACCCCACCCAAGTCCCCGGTCTACCTGGACCTGGCCTACCTGCCCA
GCGGGAGCAGCGCCCACCTGGTGGATGAGGAGTTCTTCCAGCGCGTGCGCGCGCTCTGCTACGTCATCAG
TGGCCAGGACCAGCGCAAGGAGGAAGGCATGCGGGCCGTCCTGGACGCGCTACTGGCCAGCAAGCAGCAT
TGGGACCGTGACCTGCAGGTGACCCTGATCCCCACTTTCGACTCGGTGGCCATGCATACGTGGTACGCAG
AGACGCACGCCCGGCACCAGGCGCTGGGCATCACGGTGTTGGGCAGCAACAGCATGGTGTCCATGCAGGA
TGACGCCTTCCCGGCCTGCAAGGTGGAGTTCTAGCCCCATCGCCGACACGCCCCCCACTCAGCCCAGCCC
GCCTGTCCCTAGATTCAGCCACATCAGAAATAAACTGTGACTACACTTG

Human VCY2IP1 mRNA sequence - var7 (public gi: 13623504) (SEQ ID NO: 220)
GGCACGAGGCCCTGTATGATGAGCTCCGGAACCTTCTGTTGGACCCTGCCTCTCACAAGCTACTGGTGTT
GGCTGGGCCCTGCCTGGAGGAGACGGGGGAGCTGCTGCTACAGACAGGGGGCTTCTCGCCTCACCACTTC
CTCCAGGTCCTGAAGGACAGAGAGATCCGGGACATCCTGGCCACCACGCCCCCACCTGTGCAGCCGCCCA
TACTCACCATCACCTGCCCCACCTTCGGTGACTGGGCTCAGCTGGCACCCGCTGTGCCTGGCCTTCAGGG
GGCGCTCCGGCTCCAGCTGCGGCTGAACCCCCCGGCGCAGCTGCCCAACTCTGAGGGCCTGTGCGAATTC
CTGGAGTACGTGGCTGAGTCTCTGGAGCCACCGTCCCCCTTCGAGCTGCTGGAGCCCCCGACCTCCGGGG
GCTTCCTCAGGCTGGGCCGGCCCTGCTGCTACATCTTCCCTGGAGGCCTCGGGGATGCCGCCTTCTTCGC
CGTCAATGGCTTCACTGTGCTGGTCAACGGTGGCTCAAACCCCAAGTCCAGTTTCTGGAAGCTGGTGCGG
CACCTGGACCGCGTGGATGCCGTGCTGGTGACCCACCCTGGCGCCGACAGCCTCCCCGGCCTCAACAGCC
TGCTGCGGCGCAAACTGGCGGAGCGCTCCGAGGTGGCTGCTGGTGGGGGCTCCTGGGACGACAGGCTGCG
CAGGCTCATCTCCCCCAACCTGGGGGTCGTGTTCTTCAACGCCTGCGAGGCCGCGTCGCGGCTGGCGCGC
GGCGAGGATGAGGCGGAGCTGGCGCTGAGCCTCCTGGCGCAGCTGGGCATCACGCCTCTGCCACTCAGCC
GCGGCCCCGTGCCAGCCAAACCCACCGTGCTCTTCGAGAAGATGGGCGTGGGCCGGCTGGACATGTATGT
GCTGCACCCGCCCTCCGCCGGCGCCGAGCGCACGCTGGCCTGTGTGTGCGCCCTGCTGGTGTGGCACCCC
GCCGGCCCCGGCGAGAAGGTGGTGCGCGTGCTGTTCCCCGGTTGCACCCCGCCCGCCTGCCTCCTGGACG
GCCTGGTCCGCCTGCAGCACTTGAGGTTCCTGCGAGAGCCCGTGGTGACGCCCCAGGACCTGGAGGGGCC
GGGGCGAGCCGAGAGCAAAGAGAGCGTGGGCTCCCGGGACAGCTCGAAGAGAGAGGGCCTCCTGGCCACC
CACCCTAGACCTGGCCAGGAGCGCCCTGGGGTGGCCCGCAAGGAGCCAGCACGGGCTGAGGCCCCACGCA
AGACTGAGAAAGAAGCCAAGACCCCCCGGGAGTTGAAGAAAGACCCCAAACCGAGTGTCTCCCGGACCCA
GCCGCGGGAGGTGCGCCGGGCAGCCTCTTCTGTGCCCAACCTCAAGAAGACGAATGCCCAGGCGGCACCC
AAGCCCCGCAAAGCGCCCAGCACGTCCCACTCTGGCTTCCCGCCGGTGGCAAATGGACCCCGCAGCCCGC
CCAGCCTCCGATGTGGAGAAGCCAGCCCCCCCAGTGCGGCCTGCGGCTCTCCGGCCTCCCAGCTGGTGGC
CACGCCCAGCCTGGAGCTGGGGCCGATCCCAGCCGGGGAGGAGAAGGCACTGGAGCTGCCTTTGGCCGCC
AGCTCAATCCCAAGGCCACGCACACCCTCCCCTGAGTCCCACCGGAGCCCCGCAGAGGGCAGCGAGCGGC
TGTCGCTGAGCCCACTGCGGGGCGGGGAGGCCGGGCCAGACGCCTCACCCACAGTGACCACACCCACGGT
GACCACGCCCTCACTACCCGCAGAGGTGGGCTCCCCGCACTCGACCGAGGTGGACGAGTCCCTGTCGGTG
TCCTTTGAGCAGGTGCTGCCGCCATCCGCCCCCACCAGTGAGGCTGGGCTGAGCCTCCCGCTGCGTGGCC
CCCGGGCGCGGCGCTCGGCTTCCCCACACGATGTGGACCTGTGCCTGGTGTCACCCTGTGAATTTGAGCA
TCGCAAGGCGGTGCCAATGGCACCGGCACCTGCGTCCCCCGGCAGCTCGAATGACAGCAGTGCCCGGTCA
CAGGAACGGGCAGGTGGGCTGGGGGCCGAGGAGACGCCACCCACATCGGTCAGCGAGTCCCTGCCCACCC
TGTCTGACTCGGATCCCGTGCCCCTGGCCCCCGGTGCGGCAGACTCAGACGAAGACACAGAGGGCTTTGG
AGTCCCTCGCCACGACCCTTTGCCTGACCCCCTCAAGGTCCCCCCACCACTGCCTGACCCATCCAGCATC
TGCATGGTGGACCCCGAGATGCTGCCCCCCAAGACAGCACGGCAAACGGAGAACGTCAGCCGCACCCGGA
AGCCCCTGGCCCGCCCCAACTCACGCGCTGCCGCCCCCAAAGCCACTCCAGTGGCTGCTGCCAAAACCAA
GGGGCTTGCTGGTGGGGACCGTGCCAGCCGACCACTCAGTGCCCGGAGTGAGCCCAGTGAGAAGGGAGGC
CGGGCACCCCTGTCCAGAAAGTCCTCAACCCCCAAGACTGCCACTCGAGGCCCGTCGGGGTCAGCCAGCA
GCCGGCCCGGGGTGTCAGCCACCCCACCCAAGTCCCCGGTCTACCTGGACCTGGCCTACCTGCCCAGCGG
GAGCAGCGCCCACCTGGTGGATGAGGAGTTCTTCCAGCGCGTGCGCGCGCTCTGCTACGTCATCAGTGGC
CAGGACCAGCGCAAGGAGGAAGGCATGCGGGCCGTCCTGGACGCGCTACTGGCCAGCAAGCAGCATTGGG
ACCGTGACCTGCAGGTGACCCTGATCCCCACTTTCGACTCGGTGGCCATGCATACGTGGTACGCAGAGAC
GCACGCCCGGCACCAGGCGCTGGGCATCACGGTGTTGGGCAGCAACAGCATGGTGTCCATGCAGGATGAC
GCCTTCCCGGCCTGCAAGGTGGAGTTCTAGCCCCATCGCCGACACGCCCCCCACTCAGCCCAGCCCGCCT
GTCCCTAGATTCAGCCACATCAGAAATAAACTGTGACTACACTTGAAAAAAAAAAAAAAAAA

Human VCY2IP1 mRNA sequence - var8 (public gi: 10434893) (SEQ ID NO: 221)
GAACCCCAAGTCCAGTTTCTGGAAGCTGGTGCGGCACCTGGACCGCGTGGATGCCGTGCTGGTGACCCAC
CCTGGCGCCGACAGCCTCCCCGGCCTCAACAGCCTGCTGCGGCGCAAACTGGCGGAGCGCTCCGAGGTGG
CTGCTGGTGGGGGCTCCTGGGACGACAGGCTGCGCAGGCTCATCTCCCCCAACCTGGGGGTCGTGTTCTT
CAACGCCTGCGAGGCCGCGTCGCGGCTGGCGCGCGGCGAGGATGAGGCGGAGCTGGCGCTGAGCCTCCTG
GCGCAGCTGGGCATCACGCCTCTGCCACTCAGCCGCGGCCCCGTGCCAGCCAAACCCACCGTGCTCTTCG
AGAAGATGGGCGTGGGCCGGCTGGACATGTATGTGCTGCACCCGCCCTCCGCCGGCGCCGAGCGCACGCT
GGCCTCTGTGTGCGCCCTGCTGGTGTGGCACCCCGCCGGCCCCGGCGAGAAGGTGGTGCGCGTGCTGTTC

CCCGGTTGCACCCCGCCCGCCTGCCTCCTGGACGGCCTGGTCCGCCTGCAGCACTTGAGGTTCCTGCGAG
AGCCCGTGGTGACGCCCCAGGACCTGGAGGGGCCGGGGCGAGCCGAGAGCAAAGAGAGCGTGGGCTCCCG
GGACAGCTCGAAGAGAGAGGGCCTCCTGGCCACCCACCCTAGACCTGGCCAGGAGCGCCCTGGGGTGGCC
CGCAAGGAGCCAGCACGGGCTGAGGCCCCACGCAAGACTGAGAAAGAAGCCAAGACCCCCCGGGAGTTGA
AGAAAGACCCCAAACCGAGTGTCTCCCGGACCCAGCCGCGGGAGGTGCGCCGGGCAGCCTCTTCTGTGCC
CAACCTCAAGAAGACGAATGCCCAGGCGGCACCCAAGCCCCGCAAAGCGCCCAGCACGTCCCACTCTGGC
TTCCCGCCGGTGGCAAATGGACCCCGCAGCCCGCCCAGCCTCCGATGTGGAGAAGCCAGCCCCCCCAGTG
CAGCCTGCGGCTCTCCGGCCTCCCAGCTGGTGGCCACGCCCAGCCTGGAGCTGGGGCCGATCCCAGCCGG
GGAGGAGAAGGCACTGGAGCTGCCTTTGGCCGCCAGCTCAATCCCAAGGCCACGCACACCCTCCCCTGAG
TCCCACCGGAGCCCCGCAGAGGGCAGCGAGCGGCTGTCGCTGAGCCCACTGCGGGGCGGGGAGGCCGGGC
CAGACGCCTCACCCACAGTGACCACACCCACGGTGACCACGCCCTCACTACCCGCAGAGGTGGGCTCCCC
GCACTCGACCGAGGTGGACGAGTCCCTGTCGGTGTCCTTTGAGCAGGTGCTGCCGCCATCCGCCCCCACC
AGTGAGGCTGGGCTGAGCCTCCCGCTGCGTGGCCCCCGGGCGCGGCGCTCGGCTTCCCCACACGATGTGG
ACCTGTGCCTGGTGTCACCCTGTGAATTTGAGCATCGCAAGGCGGTGCCAATGGCACCGGCACCTGCGTC
CCCCGGCAGCTCGAATGACAGCAGTGCCCGGTCACAGGAACGGGCAGGTGGGCTGGGGGCCGAGGAGACG
CCACCCACATCGGTCAGCGAGTCCCTGCCCACCCTGTCTGACTCGGATCCCGTGCCCCTGGCCCCCGGTG
CGGCAGACTCAGACGAAGACACAGAGGGCTTTGGAGTCCCTCGCCACGACCCTTTGCCTGACCCCCTCAA
GGTCCCCCCACCACTGCCTGACCCATCCAGCATCTGCATGGTGGACCCCGAGATGCTGCCCCCCAAGACA
GCACGGCAAACGGAGAACGTCAGCCGCACCCGGAAGCCCCTGGCCCGCCCCAACTCACGCGCTGCCGCCC
CCAAAGCCACTCCAGTGGCTGCTGCCAAAACCAAGGGGCTTGCTGGTGGGGACCGTGCCAGCCGACCACT
CAGTGCCCGGAGTGAGCCCAGTGAGAAGGGAGGCCGGGCACCCCTGTCCAGAAAGTCCTCAACCCCCAAG
ACTGCCACTCGAGGCCCGTCGGGGTCAGCCAGCAGCCGGCCCGGGGTGTCAGCCACCCCACCCAAGTCCC
CGGTCTACCTGGACCTGGCCTACCTGCCCAGCGGGAGCAGCGCCCACCTGGTGGATGAGGAGCTCTTCCA
GCGCGTGCGCGCGCTCTGCTACGTCATCAGTGGCCAGGACCAGCGCAAGGAGGAAGGCATGCGGGCCGTC
CTGGACGCGCTACTGGCCAGCAAGCAGCATTGGGACCGTGACCTGCAGGTGACCCTGATCCCCACTTTCG
ACTCGGTGGCCATGCATACGTGGTACGCAGAGACGCACGCCCGGCACCAGGCGCTGGGCATCACGGTGTT
GGGCAGCAACAGCATGGTGTCCATGCAGGATGACGCCTTCCCGGCTTGCAAGGTGGAGTTCTAGCCCCAT
CGCCGACACGCCCCCCACTCAGCCCAGCCCGCCTGTCCCTAGATTCAGCCACATCAGAAATAAACTGTGA
CTAC

Human VCY2IP1 mRNA sequence - var9 (public gi: 7022843) (SEQ ID NO: 222)
CATCTCCCCCAACCTGGGGGTCGTGTTCTTCAACGCCTGCGAGGCCGCGTCGCGGCTGGCGCGCGGCGAG
GATGAGGCGGAGCTGGCGCTGAGCCTCCTGGCGCAGCTGGGCATCACGCCTCTGCCACTCAGCCGCGGCC
CCGTGCCAGCCAAACCCACCGTGCTCTTCGAGAAGATGGGCGTGGGCCGGCTGGACATGTATGTGCTGCA
CCCGCCCTCCGCCGGCGCCGAGCGCACGCTGGCCTCTGTGTGCGCCCTGCTGGTGTGGCACCCCGCCGGC
CCCGGCGAGAAGGTGGTGCGCGTGCTGTTCCCCGGTTGCACCCCGCCCGCCTGCCTCCTGGACGGCCTGG
TCCGCCTGCAGCACTTGAGGTTCCTGCGAGAGCCCGTGGTGACGCCCCAGGACCTGGAGGGGCCGGGGCG
AGCCGAGAGCAAAGAGAGCGTGGGCTCCCGGGACAGCTCGAAGAGAGAGGGCCTCCTGGCCACCCACCCT
AGACCTGGCCAGGAGCGCCCTGGGGTGGCCCGCAAGGAGCCAGCACGGGCTGAGGCCCCACGCAAGACTG
AGAAAGAAGCCAAGACCCCCCGGGAGTTGAAGAAAGACCCCAAACCGAGTGTCTCCCGGACCCAGCCGCG
GGAGGTGCGCCGGGCAGCCTCTTCTGTGCCCAACCTCAAGAAGACGAATGCCCAGGCGGCACCCAAGCCC
CGCAAAGCGCCCAGCACGTCCCACTCTGGCTTCCCGCCGGTGGCAAATGGACCCCGCAGCCCGCCCAGCC
TCCGATGTGGAGAAGCCAGCCCCCCCAGTGCAGCCTGCGGCTCTCCGGCCTCCCAGCTGGTGGCCACGCC
CAGCCTGGAGCTGGGGCCGATCCCAGCCGGGGAGGAGAAGGCACTGGAGCTGCCTTTGGCCGCCAGCTCA
ATCCCAAGGCCACGCACACCCTCCCCTGAGTCCCACCGGAGCCCCGCAGAGGGCAGCGAGCGGCTGTCGC
TGAGCCCACTGCGGGGCGGGGAGGCCGGGCCAGACGCCTCACCCACAGTGACCACACCCACGGTGACCAC
GCCCTCACTACCCGCAGAGGTGGGCTCCCCGCACTCGACCGAGGTGGACGAGTCCCTGTCGGTGTCCTTT
GAGCAGGTGCTGCCGCCATCCGCCCCCACCAGTGAGGCTGGGCTGAGCCTCCCGCTGCGTGGCCCCCGGG
CGCGGCGCTCGGCTTCCCCACACGATGTGGACCTGTGCCTGGTGTCACCCTGTGAATTTGAGCATCGCAA
GGCGGTGCCAATGGCACCGGCACCTGCGTCCCCCGGCAGCTCGAATGACAGCAGTGCCCGGTCACAGGAA
CGGGCAGGTGGGCTGGGGGCCGAGGAGACGCCACCCACATCGGTCAGCGAGTCCCTGCCCACCCTGTCTG
ACTCGGATCCCGTGCCCCTGGCCCCCGGTGCGGCAGACTCAGACGAAGACACAGAGGGCTTTGGAGTCCC
TCGCCACGACCCTTTGCCTGACCCCCTCAAGGTCCCCCCACCACTGCCTGACCCATCCAGCATCTGCATG
GTGGACCCCGAGATGCTGCCCCCCAAGACAGCACGGCAAACGGAGAACGTCAGCCGCACCCGGAAGCCCC
TGGCCCGCCCCAACTCACGCGCTGCCGCCCCCAAAGCCACTCCAGTGGCTGCTGCCAAAACCAAGGGGCT
TGCTGGTGGGGACCGTGCCAGCCGACCACTCAGTGCCCGGAGTGAGCCCAGTGAGAAGGGAGGCCGGGCA
CCCCTGTCCAGAAAGTCCTCAACCCCCAAGACTGCCACTCGAGGCCCGTCGGGGTCAGCCAGCAGCCGGC
CCGGGGTGTCAGCCACCCCACCCAAGTCCCCGGTCTACCTGGACCTGGCCTACCTGCCCAGCGGGAGCAG
CGCCCACCTGGTGGATGAGGAGTTCTTCCAGCGCGTGCGCGCGCTCTGCTACGTCATCAGTGGCCAGGAC
CAGCGCAAGGAGGAAGGCATGCGGGCCGTCCTGGACGCGCTACTGGCCAGCAAGCAGCATTGGGACCGTG
ACCTGCAGGTGACCCTGATCCCCACTTTCGACTCGGTGGCCATGCATACGTGGTACGCAGAGACGCACGC
CCGGCACCAGGCGCTGGGCATCACGGTGTTGGGCAGCAACGGCATGGTGTCCATGCAGGATGACGCCTTC
CCGGCCTGCAAGGTGGAGTTCTAGCCCCATCGCCGACACGCCCCCCACTCAGCCCAGCCCGCCTGTCCCT

AGATTCAGCCACATCAGAAATAAACTGTGACTACACTTG

Human VCY2IP1 Protein sequence - var1 (public gi: 22002953) (SEQ ID NO: 315)
MAAVAGSGAAAAPSSLLLVVGSEFGSPGLLTYVLEELERGIRSWDVDPGVCNLDEQLKVFVSRHSATFSS
IVKGQRSLHHRGDNLETLVLLNPSDKSLYDELRNLLLDPASHKLLVLAGLCLEETGELLLQTGGFSPHHF
LQVLKDREIRDILATTPPPVQPPILTITCPTFGDWAQPAPAVPGLQGALRLQLRLNPPAQLPNSEGLCEF
LEYVAESLEPPSPFELLEPPTSGGFLRLGRPCCYIFPGGLGDAAFFAVNGFTVLVNGGSNPKSSFWKLVR
HLDRVDAVLVTHPGADSLPGLNSLLRRKLAERSEVAAGGGSWDDRLRRLISPNLGVVFFNACEAASRLAR
GEDEAELALSLLAQLGITPLPLSRGPVPAKPTVLFEKMGVGRLDMYVLHPPSAGAERTLASVCALLVWHP
AGPGEKVVRVLFPGCTPPACLLDGLVRLQHLRFLREPVVTPQDLEGPGRAESKESVGSRDSSKREGLLAT
HPRPGQERPGVARKEPARAEAPRKTEKEAKTPRELRKDPKPSVSRTQPREVRRAASSVPNLKKTNAQAAP
KPRKAPSTSHSGFPPVANGPRSPPSLRCGEASPPSAACGSPASQLVATPSLELGPIPAGEEKALELPLAA
SSIPRPRTPSPESHRSPAEGSERLSLSPLRGGEAGPDASPTVTTPTVTTPSLPAEVGSPHSTEVDESLSV
SFEQVLPPSAPTSEAGLSLPLRGPRARRSASPHDVDLCLVSPCEFEHRKAVPMAPAPASPGSSNDSSARS
QERAGGLGAEETPPTSVSESLPTLSDSDPVPLAPGAADSDEDTEGFGVPRHDPLPDPLKVPPPLPDPSSI
CMVDPEMLPPKTARQTENVSRTRKPLARPNSRAAAPKATPVAAAKTKGLAGGDRASRPLSARSEPSEKGG
RAPLSRKSSTPKTATRGPSGSASSRPGVSATPPKSPVYLDLAYLPSGSSAHLVDEEFFQRVRALCYVISG
QDQRKEEGMRAVLDALLASKQHWDRDLQVTLIPTFDSVAMHTWYAETHARHQALGITVLGSNSMVSMQDD
AFPACKVEF

Human VCY2IP1 Protein sequence - var2 (public gi: 21739763) (SEQ ID NO: 316)
PKMAAVAGSGAAAAPSSLLLVVGSEFGSPGLLTYVLEELERGIRSWDVDPGVCNLDEQLKVFVSRHSATF
SSIVKGQRSLHHRGDNLETLVLLNPSDKSLYDELRNLLLDPASHKLLVLAGPCLEETGELLLQTGGFSPH
HFLQVLKDREIRDILATTPPPVQPPILTITCPTFGDWAQLAPAVPGLQGALRLQLRLNPPAQLPNSEGLC
EFLEYVAESLEPPSPFELLEPPTSGGFLRLGRPCCYIFPGGLGDAAFFAVNGFTVLVNGGSNPKSSFWKL
VRHLDRVDAVLVTHPGADSLPGLNSLLRRKLAERSEVAAGGGSWDDRLRRLISPNLGVVFFNACEAASRL
ARGEDEAELALSLLAQLGITPLPLSRGPVPAKPTVLFEKMGVGRLDMYVLHPPSAGAERTLASVCALLVW
HPAGPGEKVVRVLFPGCTPPAYLLDGLVRLQHLRFLREPVVTPQDLEGPGRAESKESVGSRDSSKREGLL
ATHPRPGQERPGVARKEPARAEAPRKTEKEAKTPRELKKDPKPSVSRTQPREVRRAASSVPNLKKTNAQA
APKPRKAPSTSHSGFPPVANGPRSPPSLRCGEASPPSAACGSPASQLVATPSLELGPIPAGEEKALELPL
AASSIPRPRTPSPESHRSPAEGSERLSLSPLRGGEAGPDASPTVTTPTVTTPSLPAEVGSPHSTEVDESL
SVSFEQVLPPSAPTSEAGLSLPLRGPRARRSASPHDVDLCLVSPCEFEHRKAVPMAPAPASPGSSNDSSA
RSQERAGGLGAEETPPTSVSESLPTLSDSDPVPLAPGAADSDEDTEGFGVPRHDPLPDPLKVPPPLPDPS
SICMVDPEMLPPKTARQTENVSRTRKPLARPNSRAAAPKATPVAAAKTKGLAGGDRASRPLSARSEPSEK
GGRAPLSRKSSTPKTATRGPSGSASSRPGVSATPPKSPVYLDLAYLPSGSSAHLVDEEFFQRVRALCYVI
SGQDQRKEEGMRAVLDALLASKQHWDRDLQVTLIPTFDSVAMHTWYAETHARHQALGITVLGSNSMVSMQ
DDAFPACKVEF

Human VCY2IP1 Protein sequence - var3 (public gi: 21104446) (SEQ ID NO: 317)
MGVGRLDMYVLHPPSAGAERTLASVCALLVWHPAGPGEKVVRVLFPGCTPPACLLDGLVRLQHLRFLREP
VVTPQDLEGPGRAESKESVGSRDSSKREGLLATHPRPGQERPGVARKEPARAEAPRKTEKEAKAPRELKK
DPKPSVSRTQPREVRRAASSVPNLKKTNAQAAPKPRKAPSTSHSGFPPVANGPRSPPSLRCGEASPPSAA
CGSPASQLVATPSLELGPIPAGEEKALELPLAASSIPRPRTPSPESHRSPAEGSERLSLSPLRGGEAGPD
ASPTVTTPTVTTPSLPAEVGSPHSTEVDESLSVSFEQVLPPSAPTSEAGLSLPLRGPRARRSASPHDVDL
CLVSPCEFEHRKAVPMAPAPASPGSSNDSSARSQERAGGLGAEETPPTSVSESLPTLSDSDPVPLAPGAA
DSDEDTEGFGVPRHDPLPDPLKVPPPLPDPSSICMVDPEMLPPKTARQTENVSRTRKPLARPNSRAAAPK
ATPVAAAKTKGLAGGDRASRPLSARSEPSEKGGRAPLSRKSSTPKTATRGPSGSASSRPGVSATPPKSPV
YLDLAYLPSGSSAHLVDEEFFQRVRALCYVISGQDQRKEEGMRAVLDALLASKQHWDRDLQVTLIPTFDS
VAMHTWYAETHARHQALGITVLGSNSMVSMQDDAFPACKVEF

Human VCY2IP1 Protein sequence - var4 (public gi: 14250680) (SEQ ID NO: 318)
MGVGRLDMYVLHPPSAGAERTLASVCALLVWHPAGPGEKVVRVLFPGCTPPACLLDGLVRLQHLRFLREP
VVTPQDLEGPGRAESKESVGSRDSSKREGLLATHPRPGQERPGVARKEPARAEAPRKTEKEAKTPRELKK
DPKPSVSRTQPREVRRAASSVPNLKKTNAQAAPKPRKAPSTSHSGFPPVANGPRSPPSLRCGEASPPSAA
CGSPASQLVATPSLELGPIPAGEEKALELPLAASSIPRPRTPSPESHRSPAEGSERLSLSPLRGGEAGPD
ASPTVTTPTVTTPSLPAEVGSPHSTEVDESLSVSFEQVLPPSAPTSEAGLSLPLRGPRARRSASPHDVDL
CLVSPCEFEHRKAVPMAPAPASPGSSNDSSARSQERAGGLGAEETPPTSVSESLPTLSDSDPVPLAPGAA
DSDEDTEGFGVPRHDPLPDPLKVPPPLPDPSSICMVDPEMLPPKTARQTENVSRTRKPLARPNSRAAAPK
ATPVAAAKTKGLAGGDRASRPLSARSEPSEKGGRAPLSRKSSTPKTATRGPSGSASSRPGVSATPPKSPV
YLDLAYLPSGSSAHLVDEEFFQRVRALCYVISGQDQRKEEGMRAVLDALLASKQHWDRDLQVTLIPTFDS
VAMHTWYAETHARHQALGITVLGSNSMVSMQDDAFPACKVEF

Human VCY2IP1 Protein sequence - var5 (public gi: 13938255) (SEQ ID NO: 319)
DTDRDSSTSVSFEQVLPPSAPTSEAGLSLPLRGPRARRSASPHDVDLCLVSPCEFEHRKAVPMAPAPASP
GSSNDSSARSQERAGGLGAEETPPTSVSESLPTLSDSDPVPLAPGAADSDEDTEGFGVPRHDPLPDPLKV
PPPLPDPSSICMVDPEMLPPKTARQTENVSRTRKPLARPNSRAAAPKATPVAAAKTKGLAGGDRASRPLS
ARSEPSEKGGRAPLSRKSSTPKTATRGPSGSASSRPGVSATPPKSPVYLDLAYLPSGSSAHLVDEEFFQR
VRALCYVISGQDQRKEEGMRAVLDALLASKQHWDRDLQVTLIPTFDSVAMHTWYAETHARHQALGITVLG
SNSMVSMQDDAFPACKVEF

Human VCY2IP1 Protein sequence - var6 (public gi: 14042429) (SEQ ID NO: 320)
MAAVAGSGAAAAPSSLLLVVGSEFGSPGLLTYVLEELERGIRSWDVDPGVCNLDEQLKVFVSRHSATFSS
IVKGQRSLHHRGDNLETLVLLNPSDKSLYDELRNLLLDPASHKLLVLAGLCLEETGELLLQTGGFSPHHF
LQVLKDREIRDILATTPPPVQPPILTITCPTFGDWAQPAPAVPGLQGALRLQLRLNPPAQLPNSEGLCEF
LEYVAESLEPPSPFELLEPPTSGGFLRLGRPCCYIFPGGLGDAAFFAVNGFTVLVNGGSNPKSSFWKLVR
HLDRVDAVLVTHPGADSLPGLNSLLRRKLAERSEVAAGGGSWDDRLRRLISPNLGVVFFNACEAASRLAR
GEDEAELALSLLAQLGITPLPLSRGPVPAKPTVLFEKMGVGRLDMYVLHPPSAGAERTLASVCALLVWHP
AGPGEKVVRVLFPGCTPPACLLDGLVRLQHLRFLREPVVTPQDLEGPGRAESKESVGSRDSSKREGLLAT
HPRPGQERPGVARKEPARAEAPRKTEKEAKTPRELRKDPKPSVSRTQPREVRRAASSVPNLKKTNAQAAP
KPRKAPSTSHSGFPPVANGPRSPPSLRCGEASPPSAACGSPASQLVATPSLELGPIPAGEEKALELPLAA
SSIPRPRTPSPESHRSPAEGSERLSLSPLRGGEAGPDASPTVTTPTVTTPSLPAEVGSPHSTEVDESLSV
SFEQVLPPSAPTSEAGLSLPLRGPRARRSASPHDVDLCLVSPCEFEHRKAVPMAPAPASPGSSNDSSARS
QERAGGLGAEETPPTSVSESLPTLSDSDPVPLAPGAADSDEDTEGFGVPRHDPLPDPLKVPPPLPDPSSI
CMVDPEMLPPQDSTANGERQPHPEAPGPPQLTRCRPQSHSSGCCQNQGACWWGPCQPTTQCPE

Human VCY2IP1 Protein sequence - var7 (public gi: 13623505) (SEQ ID NO: 321)
MGVGRLDMYVLHPPSAGAERTLACVCALLVWHPAGPGEKVVRVLFPGCTPPACLLDGLVRLQHLRFLREP
VVTPQDLEGPGRAESKESVGSRDSSKREGLLATHPRPGQERPGVARKEPARAEAPRKTEKEAKTPRELKK
DPKPSVSRTQPREVRRAASSVPNLKKTNAQAAPKPRKAPSTSHSGFPPVANGPRSPPSLRCGEASPPSAA
CGSPASQLVATPSLELGPIPAGEEKALELPLAASSIPRPRTPSPESHRSPAEGSERLSLSPLRGGEAGPD
ASPTVTTPTVTTPSLPAEVGSPHSTEVDESLSVSFEQVLPPSAPTSEAGLSLPLRGPRARRSASPHDVDL
CLVSPCEFEHRKAVPMAPAPASPGSSNDSSARSQERAGGLGAEETPPTSVSESLPTLSDSDPVPLAPGAA
DSDEDTEGFGVPRHDPLPDPLKVPPPLPDPSSICMVDPEMLPPKTARQTENVSRTRKPLARPNSRAAAPK
ATPVAAAKTKGLAGGDRASRPLSARSEPSEKGGRAPLSRKSSTPKTATRGPSGSASSRPGVSATPPKSPV
YLDLAYLPSGSSAHLVDEEFFQRVRALCYVISGQDQRKEEGMRAVLDALLASKQHWDRDLQVTLIPTFDS
VAMHTWYAETHARHQALGITVLGSNSMVSMQDDAFPACKVEF

Human VCY2IP1 Protein sequence - var8 (public gi: 10434894) (SEQ ID NO: 322)
MGVGRLDMYVLHPPSAGAERTLASVCALLVWHPAGPGEKVVRVLFPGCTPPACLLDGLVRLQHLRFLREP
VVTPQDLEGPGRAESKESVGSRDSSKREGLLATHPRPGQERPGVARKEPARAEAPRKTEKEAKTPRELKK
DPKPSVSRTQPREVRRAASSVPNLKKTNAQAAPKPRKAPSTSHSGFPPVANGPRSPPSLRCGEASPPSAA
CGSPASQLVATPSLELGPIPAGEEKALELPLAASSIPRPRTPSPESHRSPAEGSERLSLSPLRGGEAGPD
ASPTVTTPTVTTPSLPAEVGSPHSTEVDESLSVSFEQVLPPSAPTSEAGLSLPLRGPRARRSASPHDVDL
CLVSPCEFEHRKAVPMAPAPASPGSSNDSSARSQERAGGLGAEETPPTSVSESLPTLSDSDPVPLAPGAA
DSDEDTEGFGVPRHDPLPDPLKVPPPLPDPSSICMVDPEMLPPKTARQTENVSRTRKPLARPNSRAAAPK
ATPVAAAKTKGLAGGDRASRPLSARSEPSEKGGRAPLSRKSSTPKTATRGPSGSASSRPGVSATPPKSPV
YLDLAYLPSGSSAHLVDEELFQRVRALCYVISGQDQRKEEGMRAVLDALLASKQHWDRDLQVTLIPTFDS
VAMHTWYAETHARHQALGITVLGSNSMVSMQDDAFPACKVEF

Human VCY2IP1 Protein sequence - var9 (public gi: 7022844) (SEQ ID NO: 323)
MGVGRLDMYVLHPPSAGAERTLASVCALLVWHPAGPGEKVVRVLFPGCTPPACLLDGLVRLQHLRFLREP
VVTPQDLEGPGRAESKESVGSRDSSKREGLLATHPRPGQERPGVARKEPARAEAPRKTEKEAKTPRELKK
DPKPSVSRTQPREVRRAASSVPNLKKTNAQAAPKPRKAPSTSHSGFPPVANGPRSPPSLRCGEASPPSAA
CGSPASQLVATPSLELGPIPAGEEKALELPLAASSIPRPRTPSPESHRSPAEGSERLSLSPLRGGEAGPD
ASPTVTTPTVTTPSLPAEVGSPHSTEVDESLSVSFEQVLPPSAPTSEAGLSLPLRGPRARRSASPHDVDL
CLVSPCEFEHRKAVPMAPAPASPGSSNDSSARSQERAGGLGAEETPPTSVSESLPTLSDSDPVPLAPGAA
DSDEDTEGFGVPRHDPLPDPLKVPPPLPDPSSICMVDPEMLPPKTARQTENVSRTRKPLARPNSRAAAPK
ATPVAAAKTKGLAGGDRASRPLSARSEPSEKGGRAPLSRKSSTPKTATRGPSGSASSRPGVSATPPKSPV
YLDLAYLPSGSSAHLVDEEFFQRVRALCYVISGQDQRKEEGMRAVLDALLASKQHWDRDLQVTLIPTFDS
VAMHTWYAETHARHQALGITVLGSNGMVSMQDDAFPACKVEF

Unigene Name: SPG20 **Unigene ID:** Hs.118087

Human SPG20 mRNA sequence - var1 (public gi: 28436884) (SEQ ID NO: 367)
AGTGTAAGGGAGTGGGAGCTGGTCCGTGCCGCGGCGGCCGCGCAGGGAGCTCTCGAGGCAACGCCGGGGC
GCCCGAGGTCTGGAAGGCGCAGAAATGGAGCAAGAGCCACAAAATGGAGAACCTGCTGAAATTAAGATCA
TCAGAGAAGCATATAAGAAGGCCTTTTTATTTGTTAACAAAGGTCTGAATACAGATGAATTAGGTCAGAA
GGAAGAAGCAAAGAACTACTATAAGCAAGGAATAGGACACCTGCTCAGAGGGATCAGCATTTCATCAAAA
GAGTCTGAACACACAGGTACTGGGTGGAATCTGCTAGACAGATGCAACAGAAAATGAAAGAAACTCTAC
AGAATGTACGCACCAGGCTGGAAATTCTAGAGAAGGGTCTTGCCACTTCTCTGCAGAATGATCTTCAGGA
GGTGCCCAAGTTATATCCAGAATTTCCACCTAAAGACATGTGTGAAAAATTACCAGAGCCTCAGTCTTTT
AGTTCAGCTCCTCAGCATGCTGAAGTAAATGGAAACACCTCAACTCCAAGTGCAGGGGCAGTTGCTGCAC
CTGCTTCTCTGTCTTTACCATCACAAAGTTGTCCAGCAGAAGCTCCTCCTGCTTATACTCCTCAAGCTGC
TGAAGGTCACTACACTGTATCCTATGGAACAGATTCTGGGGAGTTTTCATCAGTTGGAGAGGAGTTTTAT
AGGAATCATTCTCAGCCACCGCCTCTTGAGACCTTAGGGCTGGATGCAGATGAATTGATTTTGATACCAA
ATGGAGTACAGATTTTTTTTGTAAATCCTGCAGGGGAGGTTAGTGCACCTTCGTATCCTGGGTACCTTCG
AATTGTGAGGTTTTTGGATAATTCTCTCGATACGGTTCTAAACCGTCCTCCCGGGTTTCTTCAGGTTTGT
GACTGGTTATATCCTCTAGTTCCTGATAGATCTCCGGTTCTGAAATGTACTGCGGGAGCCTACATGTTTC
CTGATACAATGCTACAAGCAGCAGGATGCTTTGTGGGGGTCGTCCTGTCCTCTGAGTTACCAGAGGATGA
TAGAGAGCTCTTTGAGGATCTGTTAAGGCAAATGTCTGACCTTCGGCTCCAGGCCAACTGGAACAGAGCA
GAAGAAGAAAATGAATTCCAAATCCCTGGAAGAACTAGACCCTCCTCTGACCAACTAAAAGAAGCCTCTG
GCACTGATGTGAAACAGTTGGACCAAGGCAATAAGGATGTACGTCATAAAGGAAAACGTGGAAAAAGGGC
TAAAGATACTTCAAGTGAAGAAGTTAACCTGAGTCACATTGTACCATGTGAGCCAGTTCCAGAAGAAAAG
CCAAAAGAATTACATGAATGGAGTGAAAAAGTGGCTCACAACATTTTGTCAGGTGCTTCCTGGGTGAGTT
GGGGTTTAGTCAAAGGTGCTGAGATTACTGGTAAGGCAATCCAGAAAGGTGCTTCTAAACTCCGAGAGCG
GATTCAACCAGAAGAAAAACCCGTGGAAGTTAGTCCAGCTGTCACCAAGGGACTTTATATAGCGAAGCAA
GCTACAGGAGGAGCAGCAAAAGTCAGTCAGTTCCTGGTTGATGGAGTTTGCACTGTAGCAAATTGCGTTG
GAAAAGAACTAGCTCCACATGTCAAGAAGCATGGAAGCAAACTTGTTCCAGAATCTCTTAAAAAAGACAA
AGATGGGAAATCTCCTCTGGATGGTGCTATGGTTGTAGCAGCGAGTAGTGTTCAAGGATTTTCAACTGTC
TGGCAAGGATTGGAATGTGCAGCTAAATGCATCGTTAACAATGTTTCAGCAGAAACTGTACAAACTGTCA
GATACAAATACGGATATAATGCAGGAGAAGCTACCCACCATGCGGTGGATTCTGCGGTCAATGTTGGCGT
AACTGCCTACAATATTAACAACATTGGTATCAAAGCAATGGTGAAGAAAACTGCAACACAAACAGGACAC
ACTCTCCTTGAGGACTATCAGATAGTTGATAATTCTCAGAGGGAAAATCAAGAAGGAGCAGCAAATGTCA
ACGTGAGAGGGGAGAAGGATGAGCAGACGAAGGAAGTAAAGGAGGCAAAGAAGAAAGATAAATGATGAAG
TGCTGGGAATCACTTATACCAAAGCCTTATGAAATGGATGAAATTTTGTTAAATAGGCAAATGTGGAATT
CCTCACAGATTAACCAGTATTTTTTAAATGTATTCATTCCTACAAATTAACTTTCATAAATTTTATGGCA
TGTCTTCTATTTAAAAGGAAAAGAATAAGTATTCTTGCATCTGGCCTTAGAAATGTGAAGTTATATTCTC
AAGTTTATTTTTTTCCAAGTGTAGCTAAAATATTTTTGCAGGTAAAATAAAGCTGATAGTACATGTGTTG
TTCAAACCTTGTTAAACCTAATATTGAACTATTTTTATATCTGCTGTCTTTCAGAAGGCAAATAGGAAAC
TATATATTTGCTTAAAAATTGGCATTTAGTAACCTTAATTCTTTTTATAGAAGGAATGACTTAAAGTATT
GTCCCCTCTTTTTGCACTAATTGTGGATTTTTTTAGATGCTTCTCAAAATTTTCAGTGTGTAAGCTAAAC
AAAAACTAAAACTAAGAATTCTCAAAAAGACTTGTTCAAAACAGGGAAAGACTGATGAAAAGTAAAATGG
ACTACTTTTGTAACTTACCTGTTTGTTAGGAAATGGAATGGTTTCTTTGATTTAAAATGAATAAAAATAG
ATTATTACGTCTTTTGTATTGAGACTGTATTGTTATGAGCCTAGGAAATTTGGGAACATGATTGTATTGT
ATTAAAATTCGAAGTGATTATTATCAGCTTAATTGGATTAAAAAAGTACTTCAAGAAATTAAAAAAAAAA
AAAAAAAAAAAATAAAAAAAAAAAAAAA

Human SPG20 mRNA sequence - var2 (public gi: 7023530) (SEQ ID NO: 368)
AGGGAGCTCTCGAGGCAACGCCGGGGCGCCCGAGGTCTGGAAGGCGCAGAAATGGAGCAAGAGCCACAAA
ATGGAGAACCTGCTGAAATTAAGATCATCAGAGAAGCATATAAGAAGGCCTTTTTATTTGTTAACAAAGG
TCTGAATACAGATGAATTAGGTCAGAAGGAAGAAGCAAAGAACTACTATAAGCAAGGAATAGGACACCTG
CTCAGAGGGATCAGCATTTCATCAAAAGAGTCTGAACACACAGGTCCTGGGTGGAATCTGCTAGACAGA
TGCAACAGAAAATGAAAGAAACTCTACAGAATGTACGCACCAGGCTGGAAATTCTAGAGAAGGGTCTTGC
CACTTCTCTGCAGAATGATCTTCAGGAGGTGCCCAAGTTATATCCAGAATTTCCACCTAAAGACATGTGT
GAAAAATTACCAGAGCCTCAGTCTTTTAGTTCAGCTCCTCAGCATGCTGAAGTAAATGGAAACACCTCAA
CTCCAAGTGCAGGGGCAGTTGCTGCACCTGCTTCTCTGTCTTTACCATCACAAAGTTGTCCAGCAGAAGC
TCCTCCTGCTTATACTCCTCAAGCTGCTGAAGGTCACTACACTGTATCCTATGGAACAGATTCTGGGGAG
TTTTCATCAGTTGGAGAGGAGTTTTATAGGAATCATTCTCAGCCACCGCCTCTTGAGACCTTAGGGCTGG
ATGCAGATGAATTGATTTTGATACCAAATGGAGTACAGATTTTTTTTGTAAATCCTGCAGGGGAGGTTAG
TGCACCTTCGTATCCTGGGTACCTTCGAATTGTGAGGTTTTTGGATAATTCTCTCGATACGGTTCTAAAC
CGTCCTCCCGGGTTTCTTCAGGTTTGTGACTGGTTATATCCTCTAGTTCCTGATAGATCTCCGGTTCTGA
AATGTACTGCGGGAGCCTACATGTTTCCTGATACAATGCTACAAGCAGCAGGATGCTTTGTGGGGGTCGT
CCTGTCCTCTGAGTTACCAGAGGATGATAGAGAGCTCTTTGAGGATCTGTTAAGGCAAATGTCTGACCTT

CGGCTCCAGGCCAACTGGAACAGAGCAGAAGAAGAAAATGAATTCCAAATCCCTGGAAGAACTAGACCCT
CCTCTGACCAACTAAAAGAAGCCTCTGGCACTGATGTGAAACAGTTGGACCAAGGCAATAAGGATGTACG
TCATAAAGGAAAACGTGGAAAAAGGGCTAAAGATACTTCAAGTGAAGAAGTTAACCTGAGTCACATTGTA
CCATGTGAGCCAGTTCCAGAAGAAAAGCCAAAAGAATTACCTGAATGGAGTGAAAAAGTGGCTCACAACA
TTTTGTCAGGTATTACAGTAATGTTAATTTTTTCCCTGTATGACATTAAGCCTTTGGAACCAAATAAAG
ATATTGTTTATTAGGGAATACTGAGAAAGATAATATTTTGTATTTTGGTTTTAAATGATCAATTTAGAAA
TAAATGTAGAAGGAACTAGTCTTTTGAAACATCAGATATTGTCATATAAGTATAAATTCTTCCTGGCCTA
TTATTCTGTTTTACTATTGGGAAAATGGATAGTGAAAAGCTTCAGGAATCTTCAAATTCTTAATAGTTCT
GAATCTAAAATTAGTTATGTTCGTTTCCCCTTTGAAGCTCCCTCTTAACCTCCCCCTACCCCTGTCCCTC
AGCTGTGGTCTGAATGTGTCCCTTCAAAATTCATATATTGAAATCCTAACCCCTGAGGTGATGGTTTTAG
GAGGTGGGGCCTTTGGAAGGTGATTAGGTCATGAGGGAGGAGCCCTCATCAATGGGATTAGTCCCTTATA
AAAGAGATCCCAAAGAGCTGCCTTGTCCCTTTCACTATGTGAGGAAGCAGTAAGAAGGTGTCATTCTATG
AACCAGGAAGTGGGCCCCTCACCAGAGACCAAATGTACCAGCACCTTAGTCTTGTACTTCCCAGCCTCTA
GAATTGTGAGAAATAAATTTTTGTTGTTAAT

Human SPG20 mRNA sequence - var3 (public gi: 7023938) (SEQ ID NO: 369)
GATAATTCTCTCGATACGGTTCTAAACCGTCCTCCCGGGTTTCTTCAGGTTTGTGACTGGTTATATCCTC
TAGTTCCTGATAGATCTCCGGTTCTGAAATGTACTGCGGGAGCCTACATGTTTCCTGATACAATGCTACA
AGCAGCAGGATGCTTTGTGGGGGTCGTCCTGTCCTCTGAGTTACCAGAGGATGATAGAGAGCTCTTTGAG
GATCTGTTAAGGCAAATGTCTGACCTTCGGCTCCAGGCCAACTGGAACAGAGCAGAAGAAGAAAATGAAT
TCCAAATCCCTGGAAGAACTAGACCCTCCTCTGACCAACTAAAAGAAGCCTCTGGCACTGATGTGAAACA
GTTGGACCAAGGCAATAAGGATGTACGTCATAAAGGAAAACGTGGAAAAAGGGCTAAAGATACTTCAAGT
GAAGAAGTTAACCTGAGTCACATTGTACCATGTGAGCCAGTTCCAGAAGAAAAGCCAAAAGAATTACCTG
AACGGAGTGAAAAAGTGGCTCACAACATTTTGTCAGGTGCTTCCTGGGTGAGTTGGGGTTTAGTCAAAGG
TGCTGAGATTACTGGTAAGGCAATCCAGAAAGGTGCTTCTAAACTCCGAGAGCGGATTCAACCAGAAGAA
AAACCCGTGGAAGTTAGTCCAGCTGTCACCAAGGGACTTTATATAGCGAAGCAAGCTACAGGAGGAGCAG
CAAAAGTCAGTCAGTTCCTGGTTGATGGAGTTTGCACTGTAGCAAATTGCGTTGGAAAAGAACTAGCTCC
ACATGTCAAGAAGCATGGAAGCAAACTTGTTCCAGAATCTCTTAAAAAAGACAAAGATGGGAAATCTCCT
CTGGATGGTGCTATGGTTGTAGCAGCAAGTAGTGTTCAAGGATTTTCAACTGTCTGGCAAGGATTGGAAT
GTGCAGCTAAATGCATCGTTAACAATGTTTCAGCAGAAACTGTACAAACTGTCAGATACAAATACGGATA
TAATGCAGGAGAAGCTACCCACCATGCGGTGGATTCTGCGGTCAATGTTGGCGTAACTGCCTACAATATT
AACAACATTGGTATCAAAGCAATGGTGAAGAAAACTGCAACACAAACAGGACACACTCTCCTTGAGGACT
ATCAGATAGTTGATAATTCTCAGAGGGAAAATCAAGAAGGAGCAGCAAATGTCAACGTGAGAGGGGAGAA
GGGTGAGCAGACGAAGGAAGTAAAGGAGGCAAAGAAGAAAGATAAATGATGAAGTGCTGGGAATCACTTA
TACCAAAGCCTTATGAAATGGATGAAATTTTGTTAAATAGGCAAATGTGGAATTCCTCACAGATTAACCA
GTATTTTTTAAATGTATTCATTCCTACAAATTAACTTTCATAAATTTTATGGCATGTCTTCTATTTAAAA
GGAAAAGAATAAGTATTCTTGCATCTGGCCTTAGAAATGTGAAGTTATATTCTCAAGTTTATTTTTTTCC
AAGTGTAGCTAAAATATTTTTGCAGGTAAAATAAAGCTGATAGTACATGTGTTGTTCAAACCTTGTTAAA
CCTAATATTGAACTATTTTTATATCTGCTGTCTTTCAGAAGGCAAATAGGAAACTATATATTTGCTTAAA
AATTGGCATTTAGTAACCTTAATTCTTTTTATAGAAGGAATGACTTAAAGTATTGTCCCCTCTTTTTGCA
CTAATTGTGGATTTTTTTAGATGCTTCTCAAAATTTTCAGTGTGTAAGCTAAACAAAAACTAAAACTAAG
AATTCTCAAAAAAACTTGTTCAAAACAGGGAAAGACTGATGAAAAGTAAAATGGACTACTTTTGTAACTT
ACCTGTTTGTTAGGAAATGGAATGGTCTCTTTGATTTAAAATGAATAAAAAATAGATTATTACGTC

Human SPG20 mRNA sequence - var4 (public gi: 16553694) (SEQ ID NO: 370)
GTGCATGTTTTCTTCAGTCCTGGAAGGAAATCATAAGTGATTTGCCCCAAAAGGATTGCTGTTGAAAATG
GAGCAAGAGCCACAAAATGGAGAACCTGCTGAAATTAAGATCATCAGAGAAGCATATAAGAAGGCCTTTT
TATTTGTTAACAAAGGTCTGAATACAGATGAATTAGGTCAGAAGGAAGAAGCAAAGAACTACTATAAGCA
AGGAATAGGACACCTGCTCAGAGGGATCAGCATTTCATCAAAAGAGTCTGAACACACAGGTCCTGGGTGG
GAATCTGCTAGACAGATGCAACAGAAAATGAAAGAAACTCTACAGAATGATCTTCGTATCCTGGGTACCT
TCGAATTGTGAGGTTTTTGGATAATTCTCTCGATACGGTTCTAAACCGTCCTCCCGGGTTTCTTCAGGTT
TGTGACTGGTTATATCCTCTAGTTCCTGATAGATCTCCGGTTCTGAAATGTACTGCGGGAGCCTACATGT
TTCCTGATACAATGCTACAAGCAGCAGGATGCTTTGTGGGGGTCGTCCTGTCCTCTGAGTTACCAGAGGA
TGATAGAGAGCTCTTTGAGGATCTGTTAAGGCAAATGTCTGACCTTCGGCTCCAGGCCAACTGGAACAGA
GCAGAAGAAGAAAATGAATTCCAAATCCCTGGAAGAACTAGACCCTCCTCTGACCAACTAAAAGAAGCCT
CTGGCACTGATGTGAAACAGTTGGACCAAGGCAATAAGGATGTACGTCATAAAGGAAAACGTGGAAAAAG
GGCTAAAGATACTTCAAGTGAAGAAGTTAACCTGAGTCACATTGTACCATGTGAGCCAGTTCCAGAAGAA
AAGCCAAAAGAATTACCTGAATGGAGTGAAAAAGTGGCTCACAACATTTTGTCAGGTGCTTCCTGGGTGA
GTTGGGGTTTAGTCAAAGGTGCTGAGATTACTGGTAAGGCAATCCAGAAAGGTGCTTCTAAACTCCGAGA
GCGGATTCAACCAGAAGAAAAACCCGTGGAAGTTAGTCCAGCTGTCACCAAGGGACTTTATATAGCGAAG
CAAGCTACAGGAGGAGCAGCAAAAGTCAGTCAGTTCCTGGTTGATGGAGTTTGCACTGTAGCAAATTGCG
TTGGAAAAGAACTAGCTCCACATGTCAAGAAGCATGGAAGCAAACTTGTTCCAGAATCTCTTAAAAAAGA

CAAAGATGGGAAATCTCCTCTGGATGGTGCTATGGTTGTAGCAGCAAGTAGTGTTCAAGGATTTTCAACT
GTCTGGCAAGGATTGGAATGTGCAGCTAAATGCATCGTTAACAATGTTTCAGCAGAAACTGTACAAACTG
TCAGATACAAATACGGATAATGCAGGAGAAGCTACCCACCATGCGGTGGATTCTGCGGTCAATGTTGGCG
TAACTGCCTACAATATTGACAACATTGGTATCAAAGCAATGGTGAAGAAAACTGCAACACAAACAGGACA
CACTCTCCTTGAGGACTATCAGATAGTTGATAATTCTCAGAGGGAAAATCAAGAAGGAGCAGCAAATGTC
AACGTGAGAGGGGAGAAGGATGAGCAGACGAAGGAAGTAAAGGAGGCAAAGAAGAAAGATAAATGATGAA
GTGCTGGGAATCACTTATACCAAAGCCTTATGAAATGGATGAAATTTTGTTAAATAGGCAAATGTGGAAT
TCCTCACAGATTAACCAGTATTTTTAAATGTATTCATTCCTACAAATTAACTTTCATAAATTTTATGGC
ATGTCTTCTATTTAAAAGGAAAAGAATAAGTATTCTTGCATCTGGCCTTAGAAATGTGAAGTTATATTCT
CAAGTTTATTTTTTTCCAAGTGTAGCTAAAATATTTTTGCAGGTAAAATAAAGCTGATAGTACATGTGTT
GTTCAAACCTTGTTAAACCTAATATTGAACTATTTTTATATCTGCTGTCTTTCAGAAGGCAAATAGGAAA
CTATATATTTGCTTAAAAATTGGCATTTAGTAACCTTAATTCTTTTTATAGAAGGAATGACTTAAAGTAT
TGTCCCCTCTTTTTGCACTAATTGTGGATTTTTTTAGATGCTTCTCAAAATTTTCAGTGTGTAAGCTAAA
CAAAAACTAAAACTAAGAATTCTCAAAAAAACTTGTTCAAAACAGGGAAAGACTGATGAAAAGTAAAATG
GACTACTTTTGTAACTTACCTGTTTGTTAGGAAATGGAATGGTCTCTTTGATTTAAAATGAATAAAAATA
GATTATTACGTC

Human SPG20 mRNA sequence - var5 (public gi: 21654722) (SEQ ID NO: 371)

ATGGAGCAAGAGCCACAAAATGGAGAACCTGCTGAAATTAAGATCATCAGAGAAGCATATAAGAAGGCCT
TTTTATTTGTTAACAAAGGTCTGAATACAGATGAATTAGGTCAGAAGGAAGAAGCAAAGAACTACTATAA
GCAAGGAATAGGACACCTGCTCAGAGGGATCAGCATTTCATCAAAAGAGTCTGAACACACAGGTCCTGGG
TGGGAATCTGCTAGACAGATGCAACAGAAAATGAAAGAAACTCTACAGAATGTACGCACCAGGCTGGAAA
TTCTAGAGAAGGGTCTTGCCACTTCTCTGCAGAATGATCTTCAGGAGGTGCCCAAGTTATATCCAGAATT
TCCACCTAAAGACATGTGTGAAAAATTACCAGAGCCTCAGTCTTTTAGTTCAGCTCCTCAGCATGCTGAA
GTAAATGGAAACACCTCAACTCCAAGTGCAGGGGCAGTTGCTGCACCTGCTTCTCTGTCTTTACCATCAC
AAAGTTGTCCAGCAGAAGCTCCTCCTGCTTATACTCCTCAAGCTGCTGAAGGTCACTACACTGTATCCTA
TGGAACAGATTCTGGGGAGTTTTCATCAGTTGGAGAGGAGTTTTATAGGAATCATTCTCAGCCACCGCCT
CTTGAGACCTTAGGGCTGGATGCAGATGAATTGATTTTGATACCAAATGGAGTACAGATTTTTTTGTAA
ATCCTGCAGGGGAGGTTAGTGCACCTTCGTATCCTGGGTACCTTCGAATTGTGAGGTTTTTGGATAATTC
TCTCGATACGGTTCTAAACCGTCCTCCCGGGTTTCTTCAGGTTTGTGACTGGTTATATCCTCTAGTTCCT
GATAGATCTCCGGTTCTGAAATGTACTGCGGGAGCCTACATGTTTCCTGATACAATGCTACAAGCAGCAG
GATGCTTTGTGGGGGTCGTCCTGTCCTCTGAGTTACCAGAGGATGATAGAGAGCTCTTTGAGGATCTGTT
AAGGCAAATGTCTGACCTTCGGCTCCAGGCCAACTGGAACAGAGCAGAAGAAGAAAATGAATTCCAAATC
CCTGGAAGAACTAGACCCTCCTCTGACCAACTAAAAGAAGCCTCTGGCACTGATGTGAAACAGTTGGACC
AAGGCAATAAGGATGTACGTCATAAAGGAAAACGTGGAAAAAGGGCTAAAGATACTTCAAGTGAAGAAGT
TAACCTGAGTCACATTGTACCATGTGAGCCAGTTCCAGAAGAAAAGCCAAAAGAATTACCTGAATGGAGT
GAAAAAGTGGCTCACAACATTTTGTCAGGTGCTTCCTGGGTGAGTTGGGGTTTAGTCAAAGGTGCTGAGA
TTACTGGTAAGGCAATCCAGAAAGGTGCTTCTAAACTCCGAGAGCGGATTCAACCAGAAGAAAAACCCGT
GGAAGTTAGTCCAGCTGTCACCAAGGGACTTTATATAGCGAAGCAAGCTACAGGAGGAGCAGCAAAAGTC
AGTCAGTTCCTGGTTGATGGAGTTTGCACTGTAGCAAATTGCGTTGGAAAAGAACTAGCTCCACATGTCA
AGAAGCATGGAAGCAAACTTGTTCCAGAATCTCTTAAAAAAGACAAAGATGGGAAATCTCCTCTGGATGG
TGCTATGGTTGTAGCAGCAAGTAGTGTTCAAGGATTTTCAACTGTCTGGCAAGGATTGGAATGTGCAGCT
AAATGCATCGTTAACAATGTTTCAGCAGAAACTGTACAAACTGTCAGATACAAATACGGATATAATGCAG
GAGAAGCTACCCACCATGCGGTGGATTCTGCGGTCAATGTTGGCGTAACTGCCTACAATATTAACAACAT
TGGTATCAAAGCAATGGTGAAGAAAACTGCAACACAAACAGGACACACTCTCCTTGAGGACTATCAGATA
GTTGATAATTCTCAGAGGGAAAATCAAGAAGGAGCAGCAAATGTCAACGTGAGAGGGGAGAAGGATGAGC
AGACGAAGGAAGTAAAGGAGGCAAAGAAGAAAGATAAATGA

Human SPG20 mRNA sequence - var6 (public gi: 22074831) (SEQ ID NO: 372)

GCGGCCGCGCAGGGAGCTCTCGAGGCAACGCCGGGGCGCCCGAGGTCTGGAAGGCGCAGAAATGGAGCAA
GAGCCACAAAATGGAGAACCTGCTGAAATTAAGATCATCAGAGAAGCATATAAGAAGGCCTTTTTATTTG
TTAACAAAGGTCTGAATACAGATGAATTAGGTCAGAAGGAAGAAGCAAAGAACTACTATAAGCAAGGAAT
AGGACACCTGCTCAGAGGGATCAGCATTTCATCAAAAGAGTCTGAACACACAGGTCCTGGGTGGGAATCT
GCTAGACAGATGCAACAGAAAATGAAAGAAACTCTACAGAATGTACGCACCAGGCTGGAAATTCTAGAGA
AGGGTCTTGCCACTTCTCTGCAGAATGATCTTCAGGAGGTGCCCAAGTTATATCCAGAATTTCCACCTAA
AGACATGTGTGAAAAATTACCAGAGCCTCAGTCTTTTAGTTCAGCTCCTCAGCATGCTGAAGTAAATGGA
AACACCTCAACTCCAAGTGCAGGGGCAGTTGCTGCACCTGCTTCTCTGTCTTTACCATCACAAAGTTGTC
CAGCAGAAGCTCCTCCTGCTTATACTCCTCAAGCTGCTGAAGGTCACTACACTGTATCCTATGGAACAGA
TTCTGGGGAGTTTTCATCAGTTGGAGAGGAGTTTTATAGGAATCATTCTCAGCCACCGCCTCTTGAGACC
TTAGGGCTGGATGCAGATGAATTGATTTTGATACCAAATGGAGTACAGATTTTTTTGTAAATCCTGCAG
GGGAGGTTAGTGCACCTTCGTATCCTGGGTACCTTCGAATTGTGAGGTTTTTGGATAATTCTCTCGATAC
GGTTCTAAACCGTCCTCCCGGGTTTCTTCAGGTTTGTGACTGGTTATATCCTCTAGTTCCTGATAGATCT

CCGGTTCTGAAATGTACTGCGGGAGCCTACATGTTTCCTGATACAATGCTACAAGCAGCAGGATGCTTTG
TGGGGGTCGTCCTGTCCTCTGAGTTACCAGAGGATGATAGAGAGCTCTTTGAGGATCTGTTAAGGCAAAT
GTCTGACCTTCGGCTCCAGGCCAACTGGAACAGAGCAGAAGAAGAAAATGAATTCCAAATCCCTGGAAGA
ACTAGACCCTCCTCTGACCAACTAAAAGAAGCCTCTGGCACTGATGTGAAACAGTTGGACCAAGGCAATA
AGGATGTACGTCATAAAGGAAAACGTGGAAAAAGGGCTAAAGATACTTCAAGTGAAGAAGTTAACCTGAG
TCACATTGTACCATGTGAGCCAGTTCCAGAAGAAAAGCCAAAAGAATTACCTGAATGGAGTGAAAAAGTG
GCTCACAACATTTTGTCAGGTGCTTCCTGGGTGAGTTGGGGTTTAGTCAAAGGTGCTGAGATTACTGGTA
AGGCAATCCAGAAAGGTGCTTCTAAACTCCGAGAGCGGATTCAACCAGAAGAAAAACCCGTGGAAGTTAG
TCCAGCTGTCACCAAGGGACTTTATATAGCGAAGCAAGCTACAGGAGGAGCAGCAAAAGTCAGTCAGTTC
CTGGTTGATGGAGTTTGCACTGTAGCAAATTGCGTTGGAAAAGAACTAGCTCCACATGTCAAGAAGCATG
GAAGCAAACTTGTTCCAGAATCTCTTAAAAAAGACAAAGATGGGAAATCTCCTCTGGATGGTGCTATGGT
TGTAGCAGCAAGTAGTGTTCAAGGATTTTCAACTGTCTGGCAAGGATTGGAATGTGCAGCTAAATGCATC
GTTAACAATGTTTCAGCAGAAACTGTACAAACTGTCAGATACAAATACGGATATAATGCAGGAGAAGCTA
CCCACCATGCGGTGGATTCTGCGGTCAATGTTGGCGTAACTGCCTACAATATTAACAACATTGGTATCAA
AGCAATGGTGAAGAAAACTGCAACACAAACAGGACACACTCTCCTTGAGGACTATCAGATAGTTGATAAT
TCTCAGAGGGAAAATCAAGAAGGAGCAGCAAATGTCAACGTGAGAGGGGAGAAGGATGAGCAGACGAAGG
AAGTAAAGGAGGCAAAGAAGAAAGATAAATGATGAAGTGCTGGGAATCACTTATACCAAAGCCTTATGAA
ATGGATGAAATTTTGTTAAATAGGCAAATGTGGAATTCCTCACAGATTAACCAGTATTTTTTAAATGTAT
TCATTCCTACAAATTAACTTTCATAAATTTTATGGCATGTCTTCTATTTAAAAGGAAAAGAATAAGTATT
CTTGCATCTGGCCTTAGAAATGTGAAGTTATATTCTCAAGTTTATTTTTTTCCAAGTGTAGCTAAAATAT
TTTTGCAGGTAAAATAAAGCTGATAGTACATGTGTTGTTCAAACCTTGTTAAACCTAATATTGAACTATT
TTTATATCTGCTGTCTTTCAGAAGGCAAATAGGAAACTATATATTTGCTTAAAAATTGGCATTTAGTAAC
CTTAATTCTTTTTATAGAAGGAATGACTTAAAGTATTGTCCCCTCTTTTTGCACTAATTGTGGATTTTTT
TAGATGCTTCTCAAAATTTTCAGTGTGTAAGCTAAACAAAAACTAAAACTAAGAATTCTCAAAAAAACTT
GTTCAAAACAGGGAAAGACTGATGAAAAGTAAAATGGACTACTTTTGTAACTTACCTGTTTGTTAGGAAA
TGGAATGGTCTCTTTGATTTAAAATAAATAAAAATAGATTATTACGTCTTTTGTATTGAGACTGTATTGT
TATGAGCCTAGGAAATTTGGGAACATGATTGTATTGTATTAAAATTCGAAGTGATTATTATCAGCTTAAT
TGGATTAAAAAAGTACTTCAAGAAATTATTTTATCATATCTGCTTCTGTTTTTCCAAAAGGTTAAAACTT
GTAAAAAAAATATATATAAACAATTGAGTTTACTAATGGTAAACATTTTTATTCTGGGATTCGGTCATTG
GAATTTATATTAAAAGACAAGTTATTAAAAAGGAAAGGTTCTATTCATAATCAGGGTAAAGAATATGAAA
ACCTTAGACGTAATCCATGGTGGATAGGCATTATGGTTTCCACTTTGGCAGAAGGCAGACTATTCACAGC
CCTATTTACTTACATAGGCTAAAAAACTATGTAACTAAATACCTAATGGTATTTAATTTTTGTTTATTGA
ATTTAAGAGATTGGTATTAGTTTTCATAGCTGTAGTCCATTCTAATAATTTCTGATCTTCTAGTGGCTAC
TTAATTAGACATTATTTGAAGCTGTCTGAAGAATGCACTTTATGAATTAAAAAACTGAATTGCCTGACCT
CGTTATCACATGAGCTTATATTTTGGGAACACATAGAACTGATGGAGGCTTTTCCTAAGGCCAAGGATAA
TGTACTAGTTGTTAAAATGGAAATAAAAGTGAAGTGGTAAAT

Human SPG20 mRNA sequence - var7 (public gi: 20070809) (SEQ ID NO: 373)
GGCGGCGGCGTGCTGCGGGCTCTGTGGCGGGAGCGAGGCCGACGGGCGGGGCCGTGCGGCCGCGTGACGC
GAAGCGTTCGAGAGCGCGCGTCGTGGAACGTCTTGGTTGCCACGGCAAGCGCGCGCGCGAGGCCTTGGGA
ACCTCGGGACCGGCCCCCGGCGAGCGCAGCGGCGCCCAGTAGTCATCTTAGTGGGATTTGGGGAAGCAAC
AGGGCTGTGTGGGGTAACCTGCCACCTTTAAGTGGAAATCAGAAATGGAGCAAGAGCCACAAAATGGAGA
ACCTGCTGAAATTAAGATCATCAGAGAAGCATATAAGAAGGCCTTTTTATTTGTTAACAAAGGTCTGAAT
ACAGATGAATTAGGTCAGAAGGAAGAAGCAAAGAACTACTATAAGCAAGGAATAGGACACCTGCTCAGAG
GGATCAGCATTTCATCAAAAGAGTCTGAACACACAGGTCCTGGGTGGAATCTGCTAGACAGATGCAACA
GAAAATGAAAGAAACTCTACAGAATGTACGCACCAGGCTGGAAATTCTAGAGAAGGGTCTTGCCACTTCT
CTGCAGAATGATCTTCAGGAGGTGCCCAAGTTATATCCAGAATTTCCACCTAAAGACATGTGTGAAAAAT
TACCAGAGCCTCAGTCTTTTAGTTCAGCTCCTCAGCATGCTGAAGTAAATGGAAACACCTCAACTCCAAG
TGCAGGGGCAGTTGCTGCACCTGCTTCTCTGTCTTTACCATCACAAAGTTGTCCAGCAGAAGCTCCTCCT
GCTTATACTCCTCAAGCTGCTGAAGGTCACTACACTGTATCCTATGGAACAGATTCTGGGGAGTTTTCAT
CAGTTGGAGAGGAGTTTTATAGGAATCATTCTCAGCCACCGCCTCTTGAGACCTTAGGGCTGGATGCAGA
TGAATTGATTTTGATACCAAATGGAGTACAGATTTTTTTTGTAAATCCTGCAGGGGAGGTTAGTGCACCT
TCGTATCCTGGGTACCTTCGAATTGTGAGGTTTTTGGATAATTCTCTCGATACGGTTCTAAACCGTCCTC
CCGGGTTTCTTCAGGTTTGTGACTGGTTATATCCTCTAGTTCCTGATAGATCTCCGGTTCTGAAATGTAC
TGCGGGAGCCTACATGTTTCCTGATACAATGCTACAAGCAGCAGGATGCTTTGTGGGGGTCGTCCTGTCC
TCTGAGTTACCAGAGGATGATAGAGAGCTCTTTGAGGATCTGTTAAGGCAAATGTCTGACCTTCGGCTCC
AGGCCAACTGGAACAGAGCAGAAGAAGAAAATGAATTCCAAATCCCTGGAAGAACTAGACCCTCCTCTGA
CCAACTAAAAGAAGCCTCTGGCACTGATGTGAAACAGTTGGACCAAGGCAATAAGGATGTACGTCATAAA
GGAAAACGTGGAAAAAGGGCTAAAGATACTTCAAGTGAAGAAGTTAACCTGAGTCACATTGTACCATGTG
AGCCAGTTCCAGAAGAAAAGCCAAAAGAATTACCTGAATGGAGTGAAAAAGTGGCTCACAACATTTTGTC
AGGTGCTTCCTGGGTGAGTTGGGGTTTAGTCAAAGGTGCTGAGATTACTGGTAAGGCAATCCAGAAAGGT
GCTTCTAAACTCCGAGAGCGGATTCAACCAGAAGAAAAACCCGTGGAAGTTAGTCCAGCTGTCACCAAGG
GACTTTATATAGCGAAGCAAGCTACAGGAGGAGCAGCAAAAGTCAGTCAGTTCCTGGTTGATGGAGTTTG

CACTGTAGCAAATTGCGTTGGAAAAGAACTAGCTCCACATGTCAAGAAGCATGGAAGTCAAACTTGTTCC
AGAATCTCTTAAAAAAGACAAAGATGGGAAATCTCCTCTGGATGGTGCTATGGTTGTAGCAGCAAGTAGT
GTTCAAGGATTTTCAACTGTCTGGCAAGGATTGGAATGTGCAGCTAAATGCATCGTTAACAATGTTTCAG
CAGAAACTGTACAAACTGTCAGATACAAATACGGATATAATGCAGGAGAAGCTACCCACCATGCGGTGGA
TTCTGCGGTCAATGTTGGCGTAACTGCCTACAATATTAACAACATTGGTATCAAAGCAATGGTGAAGAAA
ACTGCAACACAAACAGGACACACTCTCCTTGAGGACTATCAGATAGTTGATAATTCTCAGAGGGAAAATC
AAGAAGGAGCAGCAAATGTCAACGTGAGAGGGGAGAAGGATGAGCAGACGAAGGAAGTAAAGGAGGCAAA
GAAGAAAGATAAATGATGAAGTGCTGGGAATCACTTATACCAAAGCCTTATGAAATGGATGAAATTTTGT
TAAATAGGCAAATGTGGAATTCCTCACAGATTAACCAGTATTTTTTAAATGTATTCATTCCTACAAATTA
ACTTTCATAAATTTTATGGCATGTCTTCTATTTAAAAGGAAAAGAATAAGTATTCTTGCATCTGGCCTTA
GAAATGTGAAGTTATATTCTCAAGTTTATTTTTTTCCAAGTGTAGCTAAAATATTTTTGCAGGTAAAATA
AAGCTGATAGTACATGTGTTGTTCAAACCTTGTTAAACCTAATATTGAACTATTTTTATATCTGCTGTCT
TTCAGAAGGCAAATAGGAAACTATATATTTGCTTAAAAATTGGCATTTAGTAACCTTAATTCTTTTTATA
GAAGGAATGACTTAAAGTATTGTCCCCTCTTTTTGCACTAATTGTGGATTTTTTTAGATGCTTCTCAAAA
TTTTCAGTGTGTAAGCTAAACAAAAACTAAAACTAAGAATTCTCAAAAAAACTTGTTCAAAACAGGGAAA
GACTGATGAAAAGTAAAATGGACTACTTTTGTAACTTACCTGTTTGTTAGGAAATGGAATGGTCTCTTTG
ATTTAAAATGAATAAAAATAGATTATTACGTCTTTTGTATTGAGACTGTATTGTTATGAGCCTAGGAAAT
TTGGGAACATGATTGTATTGTATTAAAATTCGAAGTGATTATTATCAGCTTAATTGGATTAAAAAAGTAC
TTCAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human SPG20 mRNA sequence - var8 (public gi: 3043743) (SEQ ID NO: 374)

GCGGCCGCGCAGGGAGCTCTCGAGGCAACGCCGGGGCGCCCGAGGTCTGGAAGGCGCAGAAATGGAGCAA
GAGCCACAAAATGGAGAACCTGCTGAAATTAAGATCATCAGAGAAGCATATAAGAAGGCCTTTTTATTTG
TTAACAAAGGTCTGAATACAGATGAATTAGGTCAGAAGGAAGAAGCAAAGAACTACTATAAGCAAGGAAT
AGGACACCTGCTCAGAGGGATCAGCATTTCATCAAAAGAGTCTGAACACACAGGTCCTGGGTGGAATCT
GCTAGACAGATGCAACAGAAAATGAAAGAAACTCTACAGAATGTACGCACCAGGCTGGAAATTCTAGAGA
AGGGTCTTGCCACTTCTCTGCAGAATGATCTTCAGGAGGTGCCCAAGTTATATCCAGAATTTCCACCTAA
AGACATGTGTGAAAAATTACCAGAGCCTCAGTCTTTTAGTTCAGCTCCTCAGCATGCTGAAGTAAATGGA
AACACCTCAACTCCAAGTGCAGGGGCAGTTGCTGCACCTGCTTCTCTGTCTTTACCATCACAAAGTTGTC
CAGCAGAAGCTCCTCCTGCTTATACTCCTCAAGCTGCTGAAGGTCACTACACTGTATCCTATGGAACAGA
TTCTGGGGAGTTTTCATCAGTTGGAGAGGAGTTTTATAGGAATCATTCTCAGCCACCGCCTCTTGAGACC
TTAGGGCTGGATGCAGATGAATTGATTTTGATACCAAATGGAGTACAGATTTTTTTTGTAAATCCTGCAG
GGGAGGTTAGTGCACCTTCGTATCCTGGGTACCTTCGAATTGTGAGGTTTTTGGATAATTCTCTCGATAC
GGTTCTAAACCGTCCTCCCGGGTTTCTTCAGGTTTGTGACTGGTTATATCCTCTAGTTCCTGATAGATCT
CCGGTTCTGAAATGTACTGCGGGAGCCTACATGTTTCCTGATACAATGCTACAAGCAGCAGGATGCTTTG
TGGGGGTCGTCCTGTCCTCTGAGTTACCAGAGGATGATAGAGAGCTCTTTGAGGATCTGTTAAGGCAAAT
GTCTGACCTTCGGCTCCAGGCCAACTGGAACAGAGCAGAAGAAGAAAATGAATTCCAAATCCCTGGAAGA
ACTAGACCCTCCTCTGACCAACTAAAAGAAGCCTCTGGCACTGATGTGAAACAGTTGGACCAAGGCAATA
AGGATGTACGTCATAAAGGAAAACGTGGAAAAAGGGCTAAAGATACTTCAAGTGAAGAAGTTAACCTGAG
TCACATTGTACCATGTGAGCCAGTTCCAGAAGAAAAGCCAAAAGAATTACCTGAATGGAGTGAAAAAGTG
GCTCACAACATTTTGTCAGGTGCTTCCTGGGTGAGTTGGGGTTTAGTCAAAGGTGCTGAGATTACTGGTA
AGGCAATCCAGAAAGGTGCTTCTAAACTCCGAGAGCGGATTCAACCAGAAGAAAAACCCGTGGAAGTTAG
TCCAGCTGTCACCAAGGGACTTTATATAGCGAAGCAAGCTACAGGAGGAGCAGCAAAAGTCAGTCAGTTC
CTGGTTGATGGAGTTTGCACTGTAGCAAATTGCGTTGGAAAAGAACTAGCTCCACATGTCAAGAAGCATG
GAAGCAAACTTGTTCCAGAATCTCTTAAAAAAGACAAAGATGGGAAATCTCCTCTGGATGGTGCTATGGT
TGTAGCAGCAAGTAGTGTTCAAGGATTTTCAACTGTCTGGCAAGGATTGGAATGTGCAGCTAAATGCATC
GTTAACAATGTTTCAGCAGAAACTGTACAAACTGTCAGATACAAATACGGATATAATGCAGGAGAAGCTA
CCCACCATGCGGTGGATTCTGCGGTCAATGTTGGCGTAACTGCCTACAATATTAACAACATTGGTATCAA
AGCAATGGTGAAGAAAACTGCAACACAAACAGGACACACTCTCCTTGAGGACTATCAGATAGTTGATAAT
TCTCAGAGGGAAAATCAAGAAGGAGCAGCAAATGTCAACGTGAGAGGGGAGAAGGATGAGCAGACGAAGG
AAGTAAAGGAGGCAAAGAAGAAAGATAAATGATGAAGTGCTGGGAATCACTTATACCAAAGCCTTATGAA
ATGGATGAAATTTTGTTAAATAGGCAAATGTGGAATTCCTCACAGATTAACCAGTATTTTTTAAATGTAT
TCATTCCTACAAATTAACTTTCATAAATTTTATGGCATGTCTTCTATTTAAAAGGAAAAGAATAAGTATT
CTTGCATCTGGCCTTAGAAATGTGAAGTTATATTCTCAAGTTTATTTTTTTCCAAGTGTAGCTAAAATAT
TTTTGCAGGTAAAATAAAGCTGATAGTACATGTGTTGTTCAAACCTTGTTAAACCTAATATTGAACTATT
TTTATATCTGCTGTCTTTCAGAAGGCAAATAGGAAACTATATATTTGCTTAAAAATTGGCATTTAGTAAC
CTTAATTCTTTTTATAGAAGGAATGACTTAAAGTATTGTCCCCTCTTTTTGCACTAATTGTGGATTTTTT
TAGATGCTTCTCAAAATTTTCAGTGTGTAAGCTAAACAAAAACTAAAACTAAGAATTCTCAAAAAAACTT
GTTCAAAACAGGGAAAGACTGATGAAAAGTAAAATGGACTACTTTTGTAACTTACCTGTTTGTTAGGAAA
TGGAATGGTCTCTTTGATTTAAAATGAATAAAAATAGATTATTACGTCTTTTGTATTGAGACTGTATTGT
TATGAGCCTAGGAAATTTGGGAACATGATTGTATTGTATTAAAATTCGAAGTGATTATTATCAGCTTAAT
TGGATTAAAAAAGTACTTCAAGAAATTATTTTATCATATCTGCTTCTGTTTTTCCAAAAGGTTAAAACTT
GTAAAAAAAATATATATAAACAATTGAGTTTACTAATGGTAAACATTTTTATTCTGGATTCGGTCATTG

GAATTTATATTAAAAGACAAGTTATTAAAAAGGAAAGGTTCTATTCATAATCAGGGTAAAGAATATGAAA
ACCTTAGACGTAATCCATGGTGGATAGGCATTATGGTTTCCACTTTGGCAGAAGGCAGACTATTCACAGC
CCTATTTACTTACATAGGCTAAAAAACTATGTAACTAAATACCTAATGGTATTTAATTTTTGTTTATTGA
ATTTAAGAGATTGGTATTAGTTTTCATAGCTGTAGTCCATTCTAATAATTTCTGATCTTCTAGTGGCTAC
TTAATTAGACATTATTTGAAGCTGTCTGAAGAATGCACTTTATGAATTAAAAAACTGAATTGCCTGACCT
CGTTATCACATGAGCTTATATTTTGGGAACACATAGAACTGATGGAGGCTTTTCCTAAGGCCAAGGATAA
TGTACTAGTTGTTAAAATGGAAATAAAAGTGAAGTGGTAAAT

Human SPG20 protein sequence - var1 (public gi: 28436885) (SEQ ID NO: 386)
MEQEPQNGEPAEIKIIREAYKKAFLFVNKGLNTDELGQKEEAKNYYKQGIGHLLRGISISSKESEHTGTG
WESARQMQQKMKETLQNVRTRLEILEKGLATSLQNDLQEVPKLYPEFPPKDMCEKLPEPQSFSSAPQHAE
VNGNTSTPSAGAVAAPASLSLPSQSCPAEAPPAYTPQAAEGHYTVSYGTDSGEFSSVGEEFYRNHSQPPP
LETLGLDADELILIPNGVQIFFVNPAGEVSAPSYPGYLRIVRFLDNSLDTVLNRPPGFLQVCDWLYPLVP
DRSPVLKCTAGAYMFPDTMLQAAGCFVGVVLSSELPEDDRELFEDLLRQMSDLRLQANWNRAEEENEFQI
PGRTRPSSDQLKEASGTDVKQLDQGNKDVRHKGKRGKRAKDTSSEEVNLSHIVPCEPVPEEKPKELHEWS
EKVAHNILSGASWVSWGLVKGAEITGKAIQKGASKLRERIQPEEKPVEVSPAVTKGLYIAKQATGGAAKV
SQFLVDGVCTVANCVGKELAPHVKKHGSKLVPESLKKDKDGKSPLDGAMVVAASSVQGFSTVWQGLECAA
KCIVNNVSAETVQTVRYKYGYNAGEATHHAVDSAVNVGVTAYNINNIGIKAMVKKTATQTGHTLLEDYQI
VDNSQRENQEGAANVNVRGEKDEQTKEVKEAKKKDK

Human SPG20 protein sequence - var2 (public gi: 22074832) (SEQ ID NO: 387)
MEQEPQNGEPAEIKIIREAYKKAFLFVNKGLNTDELGQKEEAKNYYKQGIGHLLRGISISSKESEHTGPG
WESARQMQQKMKETLQNVRTRLEILEKGLATSLQNDLQEVPKLYPEFPPKDMCEKLPEPQSFSSAPQHAE
VNGNTSTPSAGAVAAPASLSLPSQSCPAEAPPAYTPQAAEGHYTVSYGTDSGEFSSVGEEFYRNHSQPPP
LETLGLDADELILIPNGVQIFFVNPAGEVSAPSYPGYLRIVRFLDNSLDTVLNRPPGFLQVCDWLYPLVP
DRSPVLKCTAGAYMFPDTMLQAAGCFVGVVLSSELPEDDRELFEDLLRQMSDLRLQANWNRAEEENEFQI
PGRTRPSSDQLKEASGTDVKQLDQGNKDVRHKGKRGKRAKDTSSEEVNLSHIVPCEPVPEEKPKELPEWS
EKVAHNILSGASWVSWGLVKGAEITGKAIQKGASKLRERIQPEEKPVEVSPAVTKGLYIAKQATGGAAKV
SQFLVDGVCTVANCVGKELAPHVKKHGSKLVPESLKKDKDGKSPLDGAMVVAASSVQGFSTVWQGLECAA
KCIVNNVSAETVQTVRYKYGYNAGEATHHAVDSAVNVGVTAYNINNIGIKAMVKKTATQTGHTLLEDYQI
VDNSQRENQEGAANVNVRGEKDEQTKEVKEAKKKDK

Human SPG20 protein sequence - var3 (public gi: 3043744) (SEQ ID NO: 388)
RPRRELSRQRRGARGLEGAEMEQEPQNGEPAEIKIIREAYKKAFLFVNKGLNTDELGQKEEAKNYYKQGI
GHLLRGISISSKESEHTGPGWESARQMQQKMKETLQNVRTRLEILEKGLATSLQNDLQEVPKLYPEFPPK
DMCEKLPEPQSFSSAPQHAEVNGNTSTPSAGAVAAPASLSLPSQSCPAEAPPAYTPQAAEGHYTVSYGTD
SGEFSSVGEEFYRNHSQPPPLETLGLDADELILIPNGVQIFFVNPAGEVSAPSYPGYLRIVRFLDNSLDT
VLNRPPGFLQVCDWLYPLVPDRSPVLKCTAGAYMFPDTMLQAAGCFVGVVLSSELPEDDRELFEDLLRQM
SDLRLQANWNRAEEENEFQIPGRTRPSSDQLKEASGTDVKQLDQGNKDVRHKGKRGKRAKDTSSEEVNLS
HIVPCEPVPEEKPKELPEWSEKVAHNILSGASWVSWGLVKGAEITGKAIQKGASKLRERIQPEEKPVEVS
PAVTKGLYIAKQATGGAAKVSQFLVDGVCTVANCVGKELAPHVKKHGSKLVPESLKKDKDGKSPLDGAMV
VAASSVQGFSTVWQGLECAAKCIVNNVSAETVQTVRYKYGYNAGEATHHAVDSAVNVGVTAYNINNIGIK
AMVKKTATQTGHTLLEDYQIVDNSQRENQEGAANVNVRGEKDEQTKEVKEAKKKDK

Unigene Name: WASF1 **Unigene ID:** Hs.75850

Human WASF1 mRNA sequence - var1 (public gi: 4507912) (SEQ ID NO: 375)
CTTCTCTTGCACTTGCGGATGATGAACTGGAATAACGATGAAAGAAAGCACATCCGATCTCAACATTCAC
GTCCTGCCCTATAACCGATTAATTAATTGATCCCCAGCTAGACTAGTGTTGGAGAAATCAGCATGTTAAA
ACAACTGTTGATGATAGCTGTTGGAGTAAAGTTGCAGTGGAAGCTATGGCTGCAAAATCGTTAAAATCTT
CAAGGTGAACTGGCACAAAGGTTAATCTCAAGATGCCGCTAGTGAAAAGAAACATCGATCCTAGGCACTT
GTGCCACACAGCACTGCCTAGAGGCATTAAGAATGAACTGGAATGTGTAACCAATATTTCCTTGGCAAAT
ATAATTAGACAACTAAGTAGCCTAAGTAAATATGCTGAAGATATATTTGGAGAATTATTCAATGAAGCAC
ATAGTTTTTCCTTCAGAGTCAACTCATTGCAAGAACGTGTGGACCGTTTATCTGTTAGTGTTACACAGCT
TGATCCAAAGGAAGAAGAATTGTCTTTGCAAGATATAACAATGAGGAAAGCTTTCCGAAGTTCTACAATT
CAAGACCAGCAGCTTTTCGATCGCAAGACTTTGCCTATTCCATTACAGGAGACGTACGATGTTTGTGAAC
AGCCTCCACCTCTCAATATACTCACTCCTTATAGAGATGATGGTAAAGAAGGTCTGAAGTTTTATACCAA
TCCTTCGTATTTCTTTGATCTATGGAAAGAAAAAATGTTGCAAGATACAGAGGATAAGAGGAAGGAAAAG
AGGAAGCAGAAGCAGAAAAATCTAGATCGTCCTCATGAACCAGAAAAAGTGCCAAGAGCACCTCATGACA
GGCGGCGAGAATGGCAGAAGCTGGCCCAAGGTCCAGAGCTGGCTGAAGATGATGCTAATCTCTTACATAA
GCATATTGAAGTTGCTAATGGCCCAGCCTCTCATTTTGAAACAAGACCTCAGACATACGTGGATCATATG

GATGGATCTTACTCACTTTCTGCCTTGCCATTTAGTCAGATGAGTGAGCTTCTGACTAGAGCTGAGGAAA
GGGTATTAGTCAGACCACATGAACCACCTCCACCTCCACCAATGCATGGAGCAGGAGATGCAAAACCGAT
ACCCACCTGTATCAGTTCTGCTACAGGTTTGATAGAAAATCGCCCTCAGTCACCAGCTACAGGCAGAACA
CCTGTGTTTGTGAGCCCCACTCCCCCACCTCCTCCACCACCTCTTCCATCTGCCTTGTCAACTTCCTCAT
TAAGAGCTTCAATGACTTCAACTCCTCCCCCTCCAGTACCTCCCCCACCTCCACCTCCAGCCACTGCTTT
GCAAGCTCCAGCAGTACCACCACCTCCAGCTCCTCTTCAGATTGCCCCTGGAGTTCTTCACCCAGCTCCT
CCTCCAATTGCACCTCCTCTAGTACAGCCCTCTCCACCAGTAGCTAGAGCTGCCCCAGTATGTGAGACTG
TACCAGTTCATCCACTCCCACAAGGTGAAGTTCAGGGGCTGCCTCCACCCCCACCACCGCCTCCTCTGCC
TCCACCTGGCATTCGACCATCATCACCTGTCACAGTTACAGCTCTTGCTCATCCTCCCTCTGGGCTACAT
CCAACTCCATCTACTGCCCCAGGTCCCCATGTTCCATTAATGCCTCCATCTCCTCCATCACAAGTTATAC
CTGCTTCTGAGCCAAAGCGCCATCCATCAACCCTACCTGTAATCAGTGATGCCAGGAGTGTGCTACTGGA
AGCAATACGAAAAGGTATTCAGCTACGCAAAGTAGAAGAGCAGCGTGAACAGGAAGCTAAGCATGAACGC
ATTGAAAACGATGTTGCCACCATCCTGTCTCGCCGTATTGCTGTTGAATATAGTGATTCGGAAGATGATT
CAGAATTTGATGAAGTAGATTGGTTGGAGTAAGAAAAATGCATTGATAAATATTACAAAACTGAATGCAA
ATGTCCTTTGTGGTGCTTGTTCCTTGAAAATGTTTGGTCATTCTAGTGTTTTGCTTTCTTTTCCTTATAA
TAAATGACCCTTTTCCTCCATAACTTTTGATTTCTAAGGAAAATATTAGCATACATTTCAAACTAAATGT
TTTACAGTGGCTTATCTTTTTTTTCCCCCTGAAAAGACTAATTTGGTCAAATAAACCACTAAGTATTAAG
CATGGACAGCTGTTGTTAGAGTAGCAGATTCAGTTTTTTGATATATCTTAATTGTGTACTTTGTGAATTT
TAATTTAAAGAAAGCAACTGAAATTGAAATCTTGAGGGCAGCTGTATCTACTAATGAGCCTTATTCCATT
TCCTGATGTTTTAAAAGAAGAAACACTGCCTTGATTATACGAATACACTCAGAAAGTACATTTAGCTTGT
AGTGTTGAATTCTCTTAAAGGAATGCTTGAATTTTTTCATTATTGTTTTATTGTTTTTATATACTTGCCT
TATTTGAATGTTTAGCAGTATCCCCTTCCCACTTATATATTGTGTGATATGATTTTGCTTGCCTATAGGA
GTTAAAAACTTTTCCATGTGAAATACTCTGACTTAAACATACATGTAACTTACATAACTGTTAAGAATAA
CAGTCTGATTTAATAAATGGTTCATTTTAAAAGTT

Human WASF1 mRNA sequence - var2 (public gi: 4927209) (SEQ ID NO: 376)
ATGCCGCTAGTGAAAAGAAACATCGATCCTAGGCACTTGTGCCACACAGCACTGCCTAGAGGCATTAAGA
ATGAACTGGAATGTGTAACCAATATTTCCTTGGCAAATATAATTAGACAACTAAGTAGCCTAAGTAAATA
TGCTGAAGATATATTTGGAGAATTATTCAATGAAGCACATAGTTTTTCCTTCAGAGTCAACTCATTGCAA
GAACGTGTGGACCGTTTATCTGTTAGTGTTACACAGCTTGATCCAAAGGAAGAAGAATTGTCTTTGCAAG
ATATAACAATGAGGAAAGCTTTCCGAAGTTCTACAATTCAAGACCAGCAGCTTTTCGATCGCAAGACTTT
GCCTATTCCATTACAGGAGACGTACGATGTTTGTGAACAGCCTCCACCTCTCAATATACTCACTCCTTAT
AGAGATGATGGTAAAGAAGGTCTGAAGTTTTATACCAATCCTTCGTATTTCTTTGATCTATGGAAAGAAA
AAATGTTGCAAGATACAGAGGATAAGAGGAAGGAAAAGAGGAAGCAGAAGCAGAAAAATCTAGATCGTCC
TCATGAACCAGAAAAAGTGCCAAGAGCACCTCATGACAGGCGGCGAGAATGGCAGAAGCTGGCCCAAGGT
CCAGAGCTGGCTGAAGATGATGCTAATCTCTTACATAAGCATATTGAAGTTGCTAATGGCCCAGCCTCTC
ATTTTGAAACAAGACCTCAGACATACGTGGATCATATGGATGGATCTTACTCACTTTCTGCCTTGCCATT
TAGTCAGATGAGTGAGCTTCTGACTAGAGCTGAGGAAAGGGTATTAGTCAGACCACATGAACCACCTCCA
CCTCCACCAATGCATGGAGCAGGAGATGCAAAACCGATACCCACCTGTATCAGTTCTGCTACAGGTTTGA
TAGAAAATCGCCCTCAGTCACCAGCTACAGGCAGAACACCTGTGTTTGTGAGCCCCACTCCCCCACCTCC
TCCACCACCTCTTCCATCTGCCTTGTCAACTTCCTCATTAAGAGCTTCAATGACTTCAACTCCTCCCCCT
CCAGTACCTCCCCCACCTCCACCTCCAGCCACTGCTTTGCAAGCTCCAGCAGTACCACCACCTCCAGCTC
CTCTTCAGATTGCCCCTGGAGTTCTTCACCCAGCTCCTCCTCCAATTGCACCTCCTCTAGTACAGCCCTC
TCCACCAGTAGCTAGAGCTGCCCCAGTATGTGAGACTGTACCAGTTCATCCACTCCCACAAGGTGAAGTT
CAGGGGCTGCCTCCACCCCCACCACCGCCTCCTCTGCCTCCACCTGGCATTCGACCATCATCACCTGTCA
CAGTTACAGCTCTTGCTCATCCTCCCTCTGGGCTACATCCAACTCCATCTACTGCCCCAGGTCCCCATGT
TCCATTAATGCCTCCATCTCCTCCATCACAAGTTATACCTGCTTCTGAGCCAAAGCGCCATCCATCAACC
CTACCTGTAATCAGTGATGCCAGGAGTGTGCTACTGGAAGCAATACGAAAAGGTATTCAGCTACGCAAAG
TAGAAGAGCAGCGTGAACAGGAAGCTAAGCATGAACGCATTGAAAACGATGTTGCCACCATCCTGTCTCG
CCGTATTGCTGTTGAATATAGTGATTCGGAAGATGATTCAGAATTTGATGAAGTAGATTGGTTGGAGTAA
GAAAAATGCATTGATAAATATTACAAAACTGAATGCAAATGTCCTTTGTGGTGCTTGTTCCTTGAAAATG
TTTGGTCA

Human WASF1 protein sequence - var1 (public gi: 4507913) (SEQ ID NO: 389)
MPLVKRNIDPRHLCHTALPRGIKNELECVTNISLANIIRQLSSLSKYAEDIFGELFNEAHSFSFRVNSLQ
ERVDRLSVSVTQLDPKEEELSLQDITMRKAFRSSTIQDQQLFDRKTLPIPLQETYDVCEQPPPLNILTPY
RDDGKEGLKFYTNPSYFFDLWKEKMLQDTEDKRKEKRKQKQKNLDRPHEPEKVPRAPHDRRREWQKLAQG
PELAEDDANLLHKHIEVANGPASHFETRPQTYVDHMDGSYSLSALPFSQMSELLTRAEERVLVRPHEPPP
PPPMHGAGDAKPIPTCISSATGLIENRPQSPATGRTPVFVSPTPPPPPPPLPSALSTSSLRASMTSTPPP
PVPPPPPPPATALQAPAVPPPPAPLQIAPGVLHPAPPPIAPPLVQPSPPVARAAPVCETVPVHPLPQGEV
QGLPPPPPPPPLPPPGIRPSSPVTVTALAHPPSGLHPTPSTAPGPHVPLMPPSPPSQVIPASEPKRHPST
LPVISDARSVLLEAIRKGIQLRKVEEQREQEAKHERIENDVATILSRRIAVEYSDSEDDSEFDEVDWLE

Unigene Name: HIP-55 Unigene ID: Hs.183373

Human HIP-55 mRNA sequence - var1 (public gi: 6470260) (SEQ ID NO: 377)
ATGGCGGCGAACCTGAGCCGGAACGGGCCAGCGCTGCAAGAGGCCTACGTGCGGGTGGTCACCGAGAAGT
CCCCGACCGACTGGGCTCTCTTTACCTATGAAGGCAACAGCAATGACATCCGCGTGGCTGGCACAGGGGA
GGGTGGCCTGGAGGAGATGGTGGAGGAGCTCAACAGCGGGAAGGTGATGTACGCCTTCTGCAGAGTGAAG
GACCCCAACTCTGGACTGCCCAAATTTGTCCTCATCAACTGGACAGGCGAGGGCGTGAACGATGTGCGGA
AGGGAGCCTGTGCCAGCCACGTCAGCACCATGGCCAGCTTCCTGAAGGGGGCCCATGTGACCATCAACGC
ACGGGCCGAGGAGGATGTGGAGCCTGAGTGCATCATGGAGAAGGTGGCCAAGGCTTCAGGTGCCAACTAC
AGCTTTCACAAGGAGAGTGGCCGCTTCCAGGACGTGGGACCCCAGGCCCCAGTGGGCTCTGTGTACCAGA
AGACCAATGCCGTGTCTGAGATTAAAAGGGTTGGTAAAGACAGCTTCTGGGCCAAAGCAGAGAAGGAGGA
GGAGAACCGTCGGCTGGAGGAAAAGCGGCGGGCCGAGGAGGCACAGCGGCAGCTGGAGCAGGAGCGCCGG
GAGCGTGAGCTGCGTGAGGCTGCACGCCGGGAGCAGCGCTATCAGGAGCAGGGTGGCGAGGCCAGCCCCC
AGAGGACGTGGGAGCAGCAGCAAGAAGTGGTTTCAAGGAACCGAAATGAGCAGGAGTCTGCCGTGCACCC
GAGGGAGATTTTCAAGCAGAAGGAGAGGGCCATGTCCACCACCTCCATCTCCAGTCCTCAGCCTGGCAAG
CTGAGGAGCCCCTTCCTGCAGAAGCAGCTCACCCAACCAGAGACCCACTTTGGCAGAGAGCCAGCTGCTG
CCATCTCAAGGCCCAGGGCAGATCTCCCTGCTGAGGAGCCGGCGCCCAGCACTCCTCCATGTCTGGTGCA
GGCAGAAGAGGAGGCTGTGTATGAGGAACCTCCAGAGCAGGAGACCTTCTACGAGCAGCCCCCACTGGTG
CAGCAGCAAGGTGCCGGCTCTGAGCACATTGACCACCACATTCAGGGCCAGGGGCTCAGTGGGCAAGGGC
TCTGTGCCCGTGCCCTGTACGACTACCAGGCAGCCGACGACACAGAGATCTCCTTTGACCCCGAGAACCT
CATCACGGGCATCGAGGTGATCGACGAAGGCTGGTGGCGTGGCTATGGGCCGGATGGCCATTTTGGCATG
TTCCCTGCCAACTACGTGGAGCTCATTGAGTGAGGCTGAGGGCGGCCGCTAGACTAGTCTAGAGAAAAAA
C

Human HIP-55 mRNA sequence - var2 (public gi: 8885629) (SEQ ID NO: 378)
GAAGCTACAGCAGCGGCGCGGAGACTGCGGGGCGGGCCATGGCGGCGAACCTGAGCCGGAACGGGCCAGC
GCTGCAAGAGGCCTACGTGCGGGTGGTCACCGAGAAGTCCCCGACCGACTGGGCTCTCTTTACCTATGAA
GGCAACAGCAATGACATCCGCGTGGCTGGCACAGGGGAGGGTGGCCTGGAGGAGATGGTGGAGGAGCTCA
ACAGCGGGAAGGTGATGTACGCCTTCTGCAGAGTGAAGGACCCCAACTCTGGACTGCCCAAATTTGTCCT
CATCAACTGGACAGGCGAGGGCGTGAACGATGTGCGGAAGGGAGCCTGTGCCAGCCACGTCAGCACCATG
GCCAGCTTCCTGAAGGGGGCCCATGTGACCATCAACGCACGGGCCGAGGAGGATGTGGAGCCTGAGTGCA
TCATGGAGAAGGTGGCCAAGGCTTCAGGTGCCAACTACAGCTTTCACAAGGAGAGTGGCCGCTTCCAGGA
CGTGGGACCCCAGGCCCCAGTGGGCTCTGTGTACCAGAAGACCAATGCCGTGTCTGAGATTAAAAGGGTT
GGTAAAGACAGCTTCTGGGCCAAAGCAGAGAAGGAGGAGGAGAACCGTCGGCTGGAGGAAAAGCGGCGGG
CCGAGGAGGCACAGCGGCAGCTGGAGCAGGAGCGCCGGGAGCGTGAGCTGCGTGAGGCTGCACGCCGGGA
GCAGCGCTATCAGGAGCAGGGTGGCGAGGCCAGCCCCCAGAGGACGTGGGAGCAGCAGCAAGAAGTGGTT
TCAAGGAACCGAAATGAGCAGGAGTCTGCCGTGCACCCGAGGGAGATTTTCAAGCAGAAGGAGAGGGCCA
TGTCCACCACCTCCATCTCCAGTCCTCAGCCTGGCAAGCTGAGGAGCCCCTTCCTGCAGAAGCAGCTCAC
CCAACCAGAGACCCACTTTGGCAGAGAGCCAGCTGCTGCCATCTCAAGGCCCAGGGCAGATCTCCCTGCT
GAGGAGCCGGCGCCCAGCACTCCTCCATGTCTGGTGCAGGCAGAAGAGGAGGCTGTGTATGAGGAACCTC
CAGAGCAGGAGACCTTCTACGAGCAGCCCCCACTGGTGCAGCAGCAAGGTGCTGGCTCTGAGCACATTGA
CCACCACATTCAGGGCCAGGGGCTCAGTGGGCAAGGGCTCTGTGCCCGTGCCCTGTACGACTACCAGGCA
GCCGACGACACAGAGATCTCCTTTGACCCCGAGAACCTCATCACGGGCATCGAGGTGATCGACGAAGGCT
GGTGGCGTGGCTATGGGCCGGATGGCCATTTTGGCATGTTCCCTGCCAACTACGTGGAGCTCATTGAGTG
AGGCTGAGGGCACATCTTGCCCTTCCCCTCTCAGACATGGCTTCCTTATTGCTGGAAGAGGAGGCCTGGG
AGTTGACATTCAGCACTCTTCCAGGAATAGGACCCCCAGTGAGGATGAGGCCTCAGGGCTCCCTCCGGCT
TGGCAGACTCAGCCTGTCACCCCAAATGCAGCAATGGCCTGGTGATTCCCACACATCCTTCCTGCATCCC
CCGACCCTCCCAGACAGCTTGGCTCTTGCCCCTGACAGGATACTGAGCCAAGCCCTGCCTGTGGCCAAGC
CCTGAGTGGCCACTGCCAAGCTGCGGGGAAGGGTCCTGAGCAGGGGCATCTGGGAGGCTCTGGCTGCCTT
CTGCATTTATTTGCCTTTTTTCTTTTTCTCTTGCTTCTAAGGGGTGGTGGCCACCACTGTTTAGAATGAC
CCTTGGGAACAGTGAACGTAGAGAATTGTTTTAGCAGAGTTTGTGACCAAAGTCAGAGTGGATCATGGT
GGTTTGGCAGCAGGGAATTTGTCTTGTTGGAGCCTGCTCTGTGCTCCCCACTCCATTTCTCTGTCCCTCT
GCCTGGGCTATGGGAAGTGGGGATGCAGATGGCCAAGCTCCCACCCTGGGTATTCAAAAACGGCAGACAC
AACATGTTCCTCCACGCGGCTCAAAAAAAAAAAAAAAAAA

Human HIP-55 mRNA sequence - var3 (public gi: 8917572) (SEQ ID NO: 379)
ATGGCGGCGAACCTGAGCCGGAACGGGCCAGCGCTGCAAGAGGCCTACGTGCGGGTGGTCACCGAGAAGT
CCCCGACCGACTGGGCTCTCTTTACCTATGAAGGCAACAGCAATGACATCCGCGTGGCTGGCACAGGGGA
GGGTGGCCTGGAGGAGATGGTGGAGGAGCTCAACAGCGGGAAGGTGATGTACGCCTTCTGCAGAGTGAAG
GACCCCAACTCTGGACTGCCCAAATTTGTTCTCATCAACTGGACAGGCGAGGGCGTGAACGATGTGCGGA

AGGGAGCCTGTTCCAGCCACGTCAGCACCATGGCCAGCTTCCTGAAGGGGGCCCATGTGACCATCAACGC
ACGGGCCGAGGAGGATGTGGAGCCTGAGTGCATCATGGAGAAGGTGGCCAAGGCTTCAGGTGCCAACTAC
AGCTTTCACAAGGAGAGTGGCCGCTTCCAGGACGTGGGACCCCAGGCCCCAGTGGGCTCTGTGTACCAGA
AGACCAATGCCGTGTCTGAGATTAAAAGGGTTGGTAAAGACAGCTTCTGGGCCAAAGCAGAGAAGGAGGA
GGAGAACCGTCGGCTGGAGGAAAAGCGGCGGGCCGAGGAGGCACAGCGGCAGCTGGAGCAGGAGCGCCGG
GAGCGTGAGCTGCGTGAGGCTGCACGCCGGGAGCAGCGCTATCAGGAGCAGGGTGGCGAGGCCAGCCCCC
AGAGTACGTGGGAGCAGCAGCAAGAAGTGGTTTCAAGGAACCGAAATGAGCAGGAGTCTGCCGTGCACCC
GAGGGAGATTTTCAAGCAGAAGGAGAGGGCCATGTCCACCACCTCCATCTCCAGTCCTCAGCCTGGCAAG
CTGAGGAGCCCCTTCCTGCAGAAGCAGCTCACCCAACCAGAGACCCACTTTGGCAGAGAGCCAGCTGCTG
CCATCTCAAGGCCCAGGGCAGATCTCCCTGCTGAGGAGCCGGCGCCCAGCACTCCTCCATGTCTGGTGCA
GGCAGAAGAGGAGGCTGTGTATGAGGAACCTCCAGAGCAGGAGACCTTCTACGAGCAGCCCCCACTGGTG
CAGCAGCAAGGTGCTGGCTCTGAGCACATTGACCACCACATTCAGGGCCAGGGGCTCAGTGGGCAAGGGC
TCTGTGCCCGTGCCCTGTACGACTACCAGGCAGCCGACGACACAGAGATCTCCTTTGACCCCGAGAACCT
CATCACGGGCATCGAGGTGATCGACGAAGGCTGGTGGCGTGGCTATGGGCCGGATGGCCATTTTGGCATG
TTCCCTGCCAACTACGTGGAGCTCATTGAGTGA

Human HIP-55 mRNA sequence - var4 (public gi: 10121214) (SEQ ID NO: 380)
GGGGCGGGCCATGGCGGCGAACCTGAGCCGGAACGGGCCAGCGCTGCAAGAGGCCTACGTGCGGGTGGTC
ACCGAGAAGTCCCCGACCGACTGGGCTCTCTTTACCTATGAAGGCAACAGCAATGACATCCGCGTGGCTG
GCACAGGGGAGGGTGGCCTGGAGGAGATGGTGGAGGAGCTCAACAGCGGGAAGGTGATGTACGCCTTCTG
CAGAGTGAAGGACCCCAACTCTGGACTGCCCAAATTTGTTCTCATCAACTGGACAGGCGAGGGCGTGAAC
GATGTGCGGAAGGGAGCCTGTTCCAGCCACGTCAGCACCATGGCCAGCTTCCTGAAGGGGGCCCATGTGA
CCATCAACGCACGGGCCGAGGAGGATGTGGAGCCTGAGTGCATCATGGAGAAGGTGGCCAAGGCTTCAGG
TGCCAACTACAGCTTTCACAAGGAGAGTGGCCGCTTCCAGGACGTGGGACCCCAGGCCCCAGTGGGCTCT
GTGTACCAGAAGACCAATGCCGTGTCTGAGATTAAAAGGGTTGGTAAAGACAGCTTCTGGGCCAAAGCAG
AGAAGGAGGAGGAGAACCGTCGGCTGGAGGAAAAGCGGCGGGCCGAGGAGGCACAGCGGCAGCTGGAGCA
GGAGCGCCGGGAGCGTGAGCTGCGTGAGGCTGCACGCCGGGAGCAGCGCTATCAGGAGCAGGGTGGCGAG
GCCAGCCCCCAGAGTACGTGGGAGCAGCAGCAAGAAGTGGTTTCAAGGAACCGAAATGAGCAGGAGTCTG
CCGTGCACCCGAGGGAGATTTTCAAGCAGAAGGAGAGGGCCATGTCCACCACCTCCATCTCCAGTCCTCA
GCCTGGCAAGCTGAGGAGCCCCTTCCTGCAGAAGCAGCTCACCCAACCAGAGACCCACTTTGGCAGAGAG
CCAGCTGCTGCCATCTCAAGGCCCAGGGCAGATCTCCCTGCTGAGGAGCCGGCGCCCAGCACTCCTCCAT
GTCTGGTGCAGGCAGAAGAGGAGGCTGTGTATGAGGAACCTCCAGAGCAGGAGACCTTCTACGAGCAGCC
CCCACTGGTGCAGCAGCAAGGTGCTGGCTCTGAGCACATTGACCACCACATTCAGGGCCAGGGGCTCAGT
GGGCAAGGGCTCTGTGCCCGTGCCCTGTACGACTACCAGGCAGCCGACGACACAGAGATCTCCTTTGACC
CCGAGAACCTCATCACGGGCATCGAGGTGATCGACGAAGGCTGGTGGCGTGGCTATGGGCCGGATGGCCA
TTTTGGCATGTTCCCTGCCAACTACGTGGAGCTCATTGAGTGAGGCTGAGGGCACATCTTGCCCTTCCCC
TCTCAGACATGGCTTCCTTATTGCTGGAAGAGGAGGCCTGGGAGTTGACATTCAGCACTCTTCCAGGAAT
AGGACCCCCAGTGAGGATGAGGCCTCAGGGCTCCCTCCGGCTTGGCAGACTCAGCCTGTCACCCCAAATG
CAGCAATGGCCTGGTGATTCCCACACATCCTTCCTGCATCCCCCGACCCTCCCAGACAGCTTGGCTCTTG
CCCCTGACAGGATACTGAGCCAAGCCCTGCCTGTGGCCAAGCCCTGAGTGGCCACTGCCAAGCTGCGGGG
AAGGGTCCTGAGCAGGGGCATCTGGGAGGCTCTGGCTGCCTTCTGCATTTATTTGCCTTTTTTTCTTTTTC
TCTTGCTTCTAAGGGGTGGTGGCCACCACTGTTTAGAATGACCCTTGGGAACAGTGAACGTAGAGAATTG
TTTTTAGCAGAGTTTGTGACCAAAGTCAGAGTGGATCATGGTGGTTTGGCAGCAGGGAATTTGTCTTGTT
GGAGCCTGCTCTGTGCTCCCCACTCCATTTCTCTGTCCCTCTGCCTGGGCTATGGGAAGTGGGGATGCAG
ATGGCCAAGCTCCCACCCTGGGTATTCAAAAACGGCAGACACAACATGTTCCTCCACGCGGCTCGCTCGA
TGCCTGCAGGCCCCAGTGTGTGCCTCAACTGATTCTGACTTCAGGAAAAGTAACACAGAGTGGCCTTGGC
CTGTTGTCTTCCCCTATTTTCTGTCCCAGCTCATCCGTGTCTCTGAAGAATAAATATGCTTTTGGAAAAA
AAAAAAAAAA

Human HIP-55 mRNA sequence - var5 (public gi: 10441969) (SEQ ID NO: 381)
GACCATCAACGCACGGGCCGAGGAGGATGTGGAGCCTGAGTGCATCATGGAGAAGGTGGCCAAGGCTTCA
GGTGCCAACTACAGCTTTCACAAGGAGAGTGGCCGCTTCCAGGACGTGGGACCCCAGGCCCCAGTGGGCT
CTGTGTACCAGAAGACCAATGCCGTGTCTGAGATTAAAAGGGTTGGTAAAGACAGCTTCTGGGCCAAAGC
AGAGAAGGAGGAGGAGAACCGTCGGCTGGAGGAAAAGCGGCGGGCCGAGGAGGCACAGCGGCAGCTGGAG
CAGGAGCGCCGGGAGCGTGAGCTGCGTGAGGCTGCACGCCGGGAGCAGCGCTATCAGGAGCAGGGTGGCG
AGGCCAGCCCCCAAAGGACGTGGGAGCAGCAGCAAGAAGTGGTTTCAAGGAACCGAAATGAGCAGGAGTC
TGCCGTGCACCCGAGGGAGATTTTCAAGCAGAAGGAGAGGGCCATGTCCACCACCTCCATCTCCAGTCCT
CAGCCTGGCAAGCTGAGGAGCCCCTTCCTGCAGAAGCAGCTCACCCAACCAGAGACCCACTTTGGCAGAG
AGCCAGCTGCTGCCATCTCAAGGCCCAGGGCAGATCTCCCTGCTGAGGAGCCGGCGCCCAGCACTCCTCC
ATGTCTGGTGCAGGCAGAAGAGGAGGCTGTGTATGAGGAACCTCCAGAGCAGGAGACCTTCTACGAGCAG
CCCCCACTGGTGCAGCAGCAAGGTGCTGGCTCTGAGCACATTGACCACCACATTCAGGGCCAGGGGCTCA
GTGGGCAAGGGCTCTGTGCCCGTGCCCTGTACGACTACCAGGCAGCCGACGACACAGAGATCTCCTTTGA

CCCCGAGAACCTCATCACGGGCATCGAGGTGATCGACGAAGGCTGGTGGCGTGGCTATGGGCCGGATGGC
CATTTTGGCATGTTCCCTGCCAACTACGTGGAGCTCATTGAGTGAGGCTGAGGGCACATCTTGCCCTTCC
CCTCTCAGACATGGCTTCCTTATTGCTGGAAGAGGAGGCCTGGGAGTTGACATTCAGCACTCTTCCAGGA
ATAGGACCCCCAGTGAGGATGAGGCCTCAGGGCTCCCTCCGGCTTGGCAGACTCAGCCTGTCACCCCAAA
TGCAGCAATGGCCTGGTGATTCCCACACATCCTTCCTGCATCCCCCGACCCTCCCAGACAGCTTGGCTCT
TGCCCCTGACAGGATACTGAGCCAAGCCCTGCCTGTGGCCAAGCCCTGAGTGGCCACTGCCAAGCTGCGG
GGAAGGGTCCTGAGCAGGGGCATCTGGGAGGCTCTGGCTGCCTTCTGCATTTATTTGCCTTTTTTCTTTT
TCTCTTGCTTCTAAGGGGTGGTGGCCACCACTGTTTAGAATGACCCTTGGGAACAGTGAACGTAGAGAAT
TGTTTTTAGCAGAGTTTGTGACCAAAGTCAGAGTGGATCATGGTGGTTTGGCAGCAGGGAATTTGTCTTG
TTGGAGCCTGCTCTGTGCTCCCCACTCCATTTCTCTGTCCCTCTGCCTGGGCTATGGGAAGTGGGGATGC
AGATGGCCAAGCTCCCACCCTGGGTATTCAAAAACGGCAGACACAACATGTTCCTCCACGCGGCTCACTC
GATGCCTGCAGGCCCCAGTGTGTGCCTCAACTGATTCTGACTTCAGGAAAAGTAACACAGAGTGGCCTTG
GCCTGTTGTCTTCCCCTATTTTCTGTCCCAGCTCATCCGTGTCTCTGAAGAACAAATATGCTTTTGGACC
ACGAAAAAAAAAAAAAAAAAAAAA

Human HIP-55 mRNA sequence - var6 (public gi: 14041995) (SEQ ID NO: 382)
AGCGGCGCGGAGACTGCGGGGCGGGCCATGGCGGCGAACCTGAGCCGGAACGGGCCAGCGCTGCAAGAGG
CCTACGTGCGGGTGGTCACCGAGAAGTCCCCGACCGACTGGGCTCTCTTTACCTATGAAGGCAACAGCAA
TGACATCCGCGTGGCTGGCACAGGGGAGGGTGGCCTGGAGGAGATGGTGGAGGAGCTCAACAGCGGGAAG
GTGATGTACGCCTTCTGCAGAGTGAAGGACCCCAACTCTGGACTGCCCAAATTTGTCCTCATCAACTGGA
CAGGCGAGGGCGTGAACGATGTGCGGAAGGGAGCCTGTGCCAGCCACGTCAGCACCATGGCCAGCTTCCT
GAAGGGGGCCCATGTGACCATCAACGCACGGGCCGAGGAGGATGTGGAGCCTGAGTGCATCATGGAGAAG
GTGGCCAAGGCTTCAGGTGCCAACTACAGCTTCCACAAGGAGAGTGGCCGCTTCCAGGACGTGGGACCCC
AGGCCCCAGTGGGCTCTGTGTACCAGAAGACCAATGCCGTGTCTGAGATTAAAAGGGTTGGTAAAGACAG
CTTCTGGGCCAAAGCAGAGAAGGAGGAGGAGAACCGTCGGCTGGAGGAAAAGCGGCGGGCCGAGGAGGCA
CAGCGGCAGCTGGAGCAGGAGCGCCGGGAGCGTGAGCTGCGTGAGGCTGCACGCCGGGAGCAGCGCTATC
AGGAGCAGGGTGGCGAGGCCAGCCCCCAGAGCAGGACGTGGGAGCAGCAGCAAGAAGTGGTTTCAAGGAA
CCGAAATGAGCAGGGGTCAACATGTGCTTCCCTCCAGGAGTCTGCCGTGCACCCGAGGGAGATTTTCAAG
CAGAAGGAGAGGGCCATGTCCACCACCTCCATCTCCAGTCCTCAGCCTGGCAAGCTGAGGAGCCCCTTCC
TGCAGAAGCAGCTCACCCAACCAGAGACCCACTTTGGCAGAGAGCCAGCTGCTGCCATCTCAAGGCCCAG
GGCAGATCTCCCTGCTGAGGAGCCGGCGCCCAGCACTCCTCCATGTCTGGTGCAGGCAGAAGAGGAGGCT
GTGTATGAGGAACCTCCAGAGCAGGAGACCTTCTACGAGCAGCCCCCACTGGTGCAGCAGCAAGGTGCTG
GCTCTGAGCACATTGACCACCACATCCAGGGCCAGGGGCTCAGTGGGCAAGGGCTCTGTGCCCGTGCCCT
GTACGACTACCAGGCAGCCGACGACACAGAGATCTCCTTTGACCCCGAGAACCTCATCACGGGCATCGAG
GTGATCGACGAAGGCTGGTGGCGTGGCTATGGGCCGGATGGCCATTTTGGCATGTTCCCTGCCAACTACG
TGGAGCTCATTGAGTGAGGCTGAGGGCACATCTTGCCCTTCCCCTCTCAGACATGGCTTCCTTATTGCTG
GAAGAGGAGGCCTGGGAGTTGACATTCAGCACTCTTCCAGGAATAGGACCCCCAGTGAGGATGAGGCCTC
AGGGCTCCCTCCGGCTTGGCAGACTCAGCCTGTCACCCCAAATGCAGCAATGGCCTGGTGATTCCCACAC
ATCCTTCCTGCATCCCCCGACCCTCCCAGACAGCTTGGCTCTTGCCCCTGACAGGATACTGAGCCAAGCC
CTGCCTGTGGCCAAGCCCTGAGTGGCCACTGCCAAGCTGCGGGGAAGGGTCCTGAGCAGGGGCATCTGGG
AGGCTCTGGCTGCCTTCTGCATTTATTTGCCTTTTTTCTTTTTCTCTTGCTTCTAAGGGGTGGTGGCCAC
CACTGTTTAGAATGACCCTTGGGAACAGTGAACGTAGAGAATTGTTTTTAGCAGAGTTTGTGACCAAAGT
CAGAGTGGATCATGGTGGTTTGGCAGCAGGGAATTTGTCTTGTTGGAGCCTGCTCTGTGCTCCCCACTCC
ATTTCTCTGTCCCTCTGCCTGGGCTATGGGAAGTGGGGATGCAGATGGCCAAGCTCCCACCCTGGGTATT
CAAAAACGGCAGACACAACATGTTCCTCCACGCGGCTCACTCGATGCCTGCAGGCCCCAGTGTGTGCCTC
AACCGATTCTGACTTCAGGAAAAGTAACACAGAGTGGC

Human HIP-55 mRNA sequence - var7 (public gi: 15079722) (SEQ ID NO: 383)
GGCACGAGGGCGGAGACTGCGGGGCGGGCCATGGCGGCGAACCTGAGCCGGAACGGGCCAGCGCTGCAAG
AGGCCTACGTGCGGGTGGTCACCGAGAAGTCCCCGACCGACTGGGCTCTCTTTACCTATGAAGGCAACAG
CAATGACATCCGCGTGGCTGGCACAGGGGAGGGTGGCCTGGAGGAGATGGTGGAGGAGCTCAACAGCGGG
AAGGTGATGTACGCCTTCTGCAGAGTGAAGGACCCCAACTCTGGACTGCCCAAATTTGTCCTCATCAACT
GGACAGGCGAGGGCGTGAACGATGTGCGGAAGGGAGCCTGTGCCAGCCACGTCAGCACCATGGCCAGCTT
CCTGAAGGGGGCCCATGTGACCATCAACGCACGGGCCGAGGAGGATGTGGAGCCTGAGTGCATCATGGAG
AAGGTGGCCAAGGCTTCAGGTGCCAACTACAGCTTTCACAAGGAGAGTGGCCGCTTCCAGGACGTGGGAC
CCCAGGCCCCAGTGGGCTCTGTGTACCAGAAGACCAATGCCGTGTCTGAGATTAAAAGGGTTGGTAAAGA
CAGCTTCTGGGCCAAAGCAGAGAAGGAGGAGGAGAACCGTCGGCTGGAGGAAAAGCGGCGGGCCGAGGAG
GCACAGCGGCAGCTGGAGCAGGAGCGCCGGGAGCGTGAGCTGCGTGAGGCTGCACGCCGGGAGCAGCGCT
ATCAGGAGCAGGGTGGCGAGGCCAGCCCCCAGAGCAGGACGTGGGAGCAGCAGCAAGAAGTGGTTTCAAG
GAACCGAAATGAGCAGGAGTCTGCCGTGCACCCGAGGGAGATTTTCAAGCAGAAGGAGAGGGCCATGTCC
ACCACCTCCATCTCCAGTCCTCAGCCTGGCAAGCTGAGGAGCCCCTTCCTGCAGAAGCAGCTCACCCAAC
CAGAGACCCACTTTGGCAGAGAGCCAGCTGCTGCCATCTCAAGGCCCAGGGCAGATCTCCCTGCTGAGGA

GCCGGCGCCCAGCACTCCTCCATGTCTGGTGCAGGCAGAAGAGGAGGCTGTGTATGAGGAACCTCCAGAG
CAGGAGACCTTCTACGAGCAGCCCCCACTGGTGCAGCAGCAAGGTGCTGGCTCTGAGCACATTGACCACC
ACATTCAGGGCCAGGGGCTCAGTGGGCAAGGGCTCTGTGCCCGTGCCCTGTACGACTACCAGGCAGCCGA
CGACACAGAGATCTCCTTTGACCCCGAGAACCTCATCACGGGCATCGAGGTGATCGACGAAGGCTGGTGG
CGTGGCTATGGGCCGGATGGCCATTTTGGCATGTTCCCTGCCAACTACGTGGAGCTCATTGAGTGAGGCT
GAGGGCACATCTTGCCCTTCCCCTCTCAGACATGGCTTCCTTATTGCTGGAAGAGGAGGCCTGGGAGTTG
ACATTCAGCACTCTTCCAGGAATAGGACCCCCAGTGAGGATGAGGCCTCAGGGCTCCCTCCGGCTTGGCA
GACTCAGCCTGTCACCCCAAATGCAGCAATGGCCTGGTGATTCCCACACATCCTTCCTGCATCCCCCGAC
CCTCCCAGACAGCTTGGCTCTTGCCCCTGACAGGATACTGAGCCAAGCCCTGCCTGTGGCCAAGCCCTGA
GTGGCCACTGCCAAGCTGCGGGGAAGGGTCCTGAGCAGGGGCATCTGGGAGGCTCTGGCTGCCTTCTGCA
TTTATTTGCCTTTTTTCTTTTTCTCTTGCTTCTAAGGGGTGGTGGCCACCACTGTTTAGAATGACCCTTG
GGAACAGTGAACGTAGAGAATTGTTTTAGCAGAGTTTGTGACCAAAGTCAGAGTGGATCATGGTGGTTT
GGCAGCAGGGAATTTGTCTTGTTGGAGCCTGCTCTGTGCTCCCCACTCCATTTCTCTGTCCCTCTGCCTG
GGCTATGGGAAGTGGGGATGCAGATGGCCAAGCTCCCACCCTGGGTATTCAAAAACGGCAGACACAACAT
GTTCCTCCACGCGGCTCACTCGATGCCTGCAGGCCCCAGTGTGTGCCTCAACTGATTCTGACTTCAGGAA
AAGTAACACAGAGTGGCCTTGGCCTGTTGTCTTCCCCTATTTCTGTCCCAGCTCATCCGTGTCTCTGAA
GAACAAATATGCTTTTGGACCACGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human HIP-55 mRNA sequence - var8 (public gi: 21619482) (SEQ ID NO: 384)
CGGGCCATGGCGGCGAACCTGAGCCGGAACGGGCCAGCGCTGCAAGAGGCCTACGTGCGGGTGGTCACCG
AGAAGTCCCCGACCGACTGGGCTCTCTTTACCTATGAAGGCAACAGCAATGACATCCGCGTGGCTGGCAC
AGGGGAGGGTGGCCTGGAGGAGATGGTGGAGGAGCTCAACAGCGGGAAGGTGATGTACGCCTTCTGCAGA
GTGAAGGACCCCAACTCTGGACTGCCCAAATTTGTCCTCATCAACTGGACAGGCGAGGGCGTGAACGATG
TGCGGAAGGGAGCCTGTGCCAGCCACGTCAGCACCATGGCCAGCTTCCTGAAGGGGGCCCATGTGACCAT
CAACGCACGGGCCGAGGAGGATGTGGAGCCTGAGTGCATCATGGAGAAGGTGGCCAAGGCTTCAGGTGCC
AACTACAGCTTTCACAAGGAGAGTGGCCGCTTCCAGGACGTGGGACCCCAGGCCCCAGTGGGCTCTGTGT
ACCAGAAGACCAATGCCGTGTCTGAGATTAAAAGGGTTGGTAAAGACAGCTTCTGGGCCAAAGCAGAGAA
GGAGGAGGAGAACCGTCGGCTGGAGGAAAAGCGGCGGGCCGAGGAGGCACAGCGGCAGCTGGAGCAGGAG
CGCCGGGAGCGTGAGCTGCGTGAGGCTGCACGCCGGGAGCAGCGCTATCAGGAGCAGGGTGGCGAGGCCA
GCCCCCAGAGGACGTGGGAGCAGCAGCAAGAAGTGGTTTCAAGGAACCGAAATGAGCAGGAGTCTGCCGT
GCACCCGAGGGAGATTTTCAAGCAGAAGGAGAGGGCCATGTCCACCACCTCCATCTCCAGTCCTCAGCCT
GGCAAGCTGAGGAGCCCCTTCCTGCAGAAGCAGCTCACCCAACCAGAGACCCACTTTGGCAGAGAGCCAG
CTGCTGCCATCTCAAGGCCCAGGGCAGATCTCCCTGCTGAGGAGCCGGCGCCCAGCACTCCTCCATGTCT
GGTGCAGGCAGAAGAGGAGGCTGTGTATGAGGAACCTCCAGAGCAGGAGACCTTCTACGAGCAGCCCCCA
CTGGTGCAGCAGCAAGGTGCTGGCTCTGAGCACATTGACCACCACATTCAGGGCCAGGGGCTCAGTGGGC
AAGGGCTCTGTGCCCGTGCCCTGTACGACTACCAGGCAGCCGACGACACAGAGATCTCCTTTGACCCCGA
GAACCTCATCACGGGCATCGAGGTGATCGACGAAGGCTGGTGGCGTGGCTATGGGCCGGATGGCCATTTT
GGCATGTTCCCTGCCAACTACGTGGAGCTCATTGAGTGAGGCTGAGGGCACATCTTGCCCTTCCCCTCTC
AGACATGGCTTCCTTATTGCTGGAAGAGGAGGCCTGGGAGTTGACATTCAGCACTCTTCCAGGAATAGGA
CCCCCAGTGAGGATGAGGCCTCAGGGCTCCCTCCGGCTTGGCAGACTCAGCCTGTCACCCCAAATGCAGC
AATGGCCTGGTGATTCCCACACATCCTTCCTGCATCCCCCGACCCTCCCAGACAGCTTGGCTCTTGCCCC
TGACAGGATACTGAGCCAAGCCCTGCCTGTGGCCAAGCCCTGAGTGGCCACTGCCAAGCTGCGGGGAAGG
GTCCTGAGCAGGGGCATTTGGGAGGCTCTGGCTGCCTTCTGCATTTATTTGCCTTTTTTCTTTTTCTCTT
GCTTCTAAGGGGTGGTGGCCACCACTGTTTAGAATGACCCTTGGGAACAGTGAACGTAGAGAATTGTTTT
TAGCAGAGTTTGTGACCAAAGTCAGAGTGGATCATGGTGGTTTGGCAGCAGGGAATTTGTCTTGTTGGAG
CCTGCTCTGTGCTCCCCACTCCATTTTCTGTCCCTCTGCCTGGGCTGTGGGAAGTGGGGATGCAGATGG
CCAAGCTCCCCCCCTGGGTATTCAAAAACGGCAGACACAACATGTTCCTCCACGCGGCTCACTCGATGCC
TGCAGGCCCCAGTGTGTGCCTCAACTGATTCTGACTTCAGGAAAAGTAACACAGAGTGGCCTTGGCCTGT
TGTCTTCCCCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Human HIP-55 mRNA sequence - var9 (public gi: 23959038) (SEQ ID NO: 385)
GGCACGAGGATTTGACACATGAATGTATAGCAGTCATTGGGAAACTCCACAGCTCATGTTTTCCTCATAG
TAGATGTGTGCTCCCATCTCCATGGCTTTGTCCCTCACAACCCCCACCCCATGGTAAGTCAGGCCAGTGT
CCTCCCAGCTGCAGAGCTGAGAAGGCTGCACAGTTGCCTACTGAGAACCTGCCTAGTGGGTCAGAGCAAA
GTGAGAACGGGCCTGTGCCCACCCACAGTGTTACTGTCAGCCCAAGCTCTTTGGGATGTAGTGGAAAGTC
ATGGTGGATACGGTGAGGAGAGATGGAACCCAAGGTGCTGGCTACAGAGCTCACTTGTGTTTCGTTTCAG
GGCTCTCTTTACCTATGAAGGCAACAGCAATGACATCCGCGTGGCTGGCACAGGGGGTGAGTATGACTCC
AAATGGACTCAGGGACACCAGGAGGTAGGAGGGTGACGACGAGGGGTCAGCGCACTCAGCTGTCTTGGTC
CACTGAGCCACATGGGGCTTCCAGTGCTCACTGGCCACTTCTGGCAGGCCTTAGGTTTCAGATATGTGTA
AGTGAAAACATTCCTCCTTGGTTCTCCTTCCCTCTGGGTGCAGGGGAGTGCTTTCTCTTTGTCTACTTGG
GGAGAGCTGAGAGGGAAACAGGCCTCTCCAGCTTGTGGGCAGCCTGCGTTGGGAGCTGCGGTGGGAAGCT
CACCAGTCCCAGAACTGGTGCTGGTGGAAAGAAAGTCCACAGACATATCTTCTTCTCCCTTTGTCCTGCC

CTGGTCTTGTGCCGAGTGCTTGCAGGGGCCCCATCCTCACTGGGAGAGGCAGTATCACTGCAGATAGTCA
CGGGGGAGGCTCTGGAGGTCTCTACAGGAAGGACAGGCTCTTGGCCAGCACAGAGCAGAGGTTGTCAGGG
TAGGCTTCGTCAGAGTGTGACCTGTGGGCCCCTCAGCTGACACCCGTGACTGCTCCTCCTCCAGAAGTTG
CCTGACCCCTCCCTCTGTCCTGAGCTGGACATGGCTTCATTGTTCAATGAACACTCGGAGTGGTTCTCCA
CGGTTTGATGTCTGTTGTTGGTAGAAAGCCCCTTCCTTTCACAATCTTTCTGGGAGGTGTCCCCTTTCTA
GAAGGATTGCCATTGAACAGTAGACATGTGGTGTGGCAGGTGACTGGGAGTTGCAGAGATCAACAACTTG
AGAGTTTCCTGTCATCCCCAGTGGCACAGGACAGGGCTCTGCCACAAATGCAACAATTTGCTGTCCCCAG
AGTGGGGCTCATGACTGCCTCCACTCATACGGAGCCCTGTAGATGAAATACCTGATCAGCTCTTCCTCCT
TATAACCTGGAAAAGTTTGTGAGGGCTAAGCCTCAGTGTCAGGGAGAATTGTTTAGAGCTGCCCACTCCT
GTGCTCCCCCTGTCCCCATCACCCTCTCTTCTGGAGTCTGAGGACTGAGCCAGTTACGCCACTGCAGGAT
GTTCAATCTGGTCTGGCCGTCTGGGTGGCCCTGGAACTTGAGCAGACACAGGTGCAGGCAGTGGTGACTC
TACAGGCCCTGCTATTCCGGGCCCTTTTGCAACGTTGTGGCAACAATAAAATTTTGACGTAGCCATCCTC
CATTTGGAAGTCTGGTGGCTGGTTTGCCGTGGAAATGACCCTGTTTTTATTTCCAGAATTACCTCTGGGT
TTAGAGAAGTGGTTTTTAAACGAGTGTGGGTAAAAAAATTACCTGAGGTACTTGTCAGAATCGCAGACTT
CTAGGTCCCACCCAGCTCTCATCAATCAGTTTAGTGAGGGTGGTGCCCAGGACTCTGATTTTAAACATAC
CCCTAGAAAGATTCTGATACAGGTAGAGGTGAGAAGCCCTGGTTTAGAAGCAGCTCGGCCTCCCTTCATG
GTGGGACCAGGGCCAGCAGGGAATGTCAGGGCCACCCCTGACCTTCACTGTGACTCTGCTGCAGAGGGTG
GCCTGGAGGAGATGGTGGAGGAGCTCAACAGCGGGAAGGTGATGTACGCCTTCTGCAGAGTGAAGGACCC
CAACTCTGGACTGCCCAAATTTGTCCTCATCAACTGGACAGGCGAGGGCGTGAACGATGTGCGGAAGGGA
GCCTGTGCCAGCCACGTCAGCACCATGGCCAGCTTCCTGAAGGGGGCCCATGTGACCATCAACGCACGGG
CCGAGGAGGATGTGGAGCCTGAGTGCATCATGGAGAAGGTGGCCAAGGCTTCAGGTGCCAACTACAGCTT
TCACAAGGAGAGTGGCCGCTTCCAGGACGTGGGACCCCAGGCCCCAGTGGGCTCTGTGTACCAGAAGACC
AATGCCGTGTCTGAGATTAAAAGGGTTGGTAAAGACAGCTTCTGGGCCAAAGCAGAGGTGAGTGCTGCCC
CGGGGCATGCTGGGCACGTGGGAGTGTTCTGCTTGCTGTGGCTCATCTTTCCTCACAAGTGAGCTCATGC
AGCATCCACTCTCCTTGGTGCCCATTACAGATGGTCACACTGAGGCTCGGGTAAGTTAAGCCACAAGGCT
AATGATCGACTGGCTCTGGTGCCCGTCTTTGGCCATGTGCCTAAAACTCAGTCCTGGGCAGGGGATTAGG
CTGAAGTGGCAGCATAGGGCTGAGCGGGCAGTGGCTCTCCCTGCAGAAGGAGGAGGAGAACCGTCGGCTG
GAGGAAAAGCGGCGGGCCGAGGAGGCACAGCGGCAGCTGGAGCAGGAGCGCCGGGAGCGTGAGCTGCGTG
AGGCTGCACGCCGGGAGCAGCGCTATCAGGAGCAGGGTGGCGAGGCCAGCCCCCAGAGGACGTGGGAGCA
GCAGCAAGAAGTGGTTTCAAGGAACCGAAATGAGCAGGAGTCTGCCGTGCACCCGAGGGAGATTTTCAAG
CAGAAGGAGAGGGCCATGTCCACCACCTCCATCTCCAGTCCTCAGCCTGGCAAGCTGAGGAGCCCCTTCC
TGCAGAAGCAGCTCACCCAACCAGAGACCCACTTTGGCAGAGAGCCAGCTGCTGCCATCTCAAGGCCCAG
GGCAGATCTCCCTGCTGAGGAGCCGGCGCCCAGCACTCCTCCATGTCTGGTGCAGGCAGAAGAGGAGGCT
GTGTATGAGGAACCTCCAGAGCAGGAGACCTTCTACGAGCAGCCCCCACTGGTGCAGCAGCAAGGTGCTG
GCTCTGAGCACATTGACCACCACATTCAGGGCCAGGGGCTCAGTGGGCAAGGGCTCTGTGCCCGTGCCCT
GTACGACTACCAGGCAGCCGACGACACAGAGATCTCCTTTGACCCCGAGAACCTCATCACGGGCATCGAG
GTGATCGACGAAGGCTGGTGGCGTGGCTATGGGCCGGATGGCCATTTTGGCATGTTCCCTGCCAACTACG
TGGAGCTCATTGAGTGAGGCTGAGGGCACATCTTGCCCTTCCCCTCTCAGACATGGCTTCCTTATTGCTG
GAAGAGGAGGCCTGGGAGTTGACATTCAGCACTCTTCCAGGAATAGGACCCCCAGTGAGGATGAGGCCTC
AGGGCTCCCTCCGGCTTGGCAGACTCAGCCTGTCACCCCAAATGCAGCAATGGCCTGGTGATTCCCACAC
ATCCTTCCTGCATCCCCCGACCCTCCCAGACAGCTTGGCTCTTGCCCCTGACAGGATACTGAGCCAAGCC
CTGCCTGTGGCCAAGCCCTGAGTGGCCACTGCCAAGCTGCGGGGAAGGGTCCTGAGCAGGGGCATCTGGG
AGGCTCTGGCTGCCTTCTGCATTTATTTGCCTTTTTTCTTTTTCTCTTGCTTCTAAGGGGTGGTGGCCAC
CACTGTTTAGAATGACCCTTGGGAACAGTGAACGTAGAGAATTGTTTTTAGCAGAGTTTGTGACCAAAGT
CAGAGTGGATCATGGTGGTTTGGCAGCAGGGAATTTGTCTTGTTGGAGCCTGCTCTGTGCTCCCCACTCC
ATTTCTCTGTCCCTCTGCCTGGGCTATGGGAAGTGGGGATGCAGATGGCCAAGCTCCCACCCTGGGTATT
CAAAAACGGCAGACACAACATGTTCCTCCACGCGGCTCACTCGATGCCTGCAGGCCCCAGTGTGTGCCTC
AACTGATTCTGACTTCAGGAAAAGTAACACAGAGTGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAA

Human HIP-55 protein sequence - var1 (public gi: 21619483) (SEQ ID NO: 390)
MAANLSRNGPALQEAYVRVVTEKSPTDWALFTYEGNSNDIRVAGTGEGGLEEMVEELNSGKVMYAFCRVK
DPNSGLPKFVLINWTGEGVNDVRKGACASHVSTMASFLKGAHVTINARAEEDVEPECIMEKVAKASGANY
SFHKESGRFQDVGPQAPVGSVYQKTNAVSEIKRVGKDSFWAKAEKEEENRRLEEKRRAEEAQRQLEQERR
ERELREAARREQRYQEQGGEASPQRTWEQQQEVVSRNRNEQESAVHPREIFKQKERAMSTTSISSPQPGK
LRSPFLQKQLTQPETHFGREPAAAISRPRADLPAEEPAPSTPPCLVQAEEEAVYEEPPEQETFYEQPPLV
QQQGAGSEHIDHHIQGQGLSGQGLCARALYDYQAADDTEISFDPENLITGIEVIDEGWWRGYGPDGHFGM
FPANYVELIE

Human HIP-55 protein sequence - var2 (public gi: 15079723) (SEQ ID NO: 391)
MAANLSRNGPALQEAYVRVVTEKSPTDWALFTYEGNSNDIRVAGTGEGGLEEMVEELNSGKVMYAFCRVK
DPNSGLPKFVLINWTGEGVNDVRKGACASHVSTMASFLKGAHVTINARAEEDVEPECIMEKVAKASGANY

SFHKESGRFQDVGPQAPVGSVYQKTNAVSEIKRVGKDSFWAKAEKEEENRRLEEKRRAEEAQRQLEQERR
ERELREAARREQRYQEQGGEASPQSRTWEQQQEVVSRNRNEQESAVHPREIFKQKERAMSTTSISSPQPG
KLRSPFLQKQLTQPETHFGREPAAAISRPRADLPAEEPAPSTPPCLVQAEEEAVYEEPPEQETFYEQPPL
VQQQGAGSEHIDHHIQGQGLSGQGLCARALYDYQAADDTEISFDPENLITGIEVIDEGWWRGYGPDGHFG
MFPANYVELIE

Human HIP-55 protein sequence - var3 (public gi: 14041996) (SEQ ID NO: 392)
MAANLSRNGPALQEAYVRVVTEKSPTDWALFTYEGNSNDIRVAGTGEGGLEEMVEELNSGKVMYAFCRVK
DPNSGLPKFVLINWTGEGVNDVRKGACASHVSTMASFLKGAHVTINARAEEDVEPECIMEKVAKASGANY
SFHKESGRFQDVGPQAPVGSVYQKTNAVSEIKRVGKDSFWAKAEKEEENRRLEEKRRAEEAQRQLEQERR
ERELREAARREQRYQEQGGEASPQSRTWEQQQEVVSRNRNEQGSTCASLQESAVHPREIFKQKERAMSTT
SISSPQPGKLRSPFLQKQLTQPETHFGREPAAAISRPRADLPAEEPAPSTPPCLVQAEEEAVYEEPPEQE
TFYEQPPLVQQQGAGSEHIDHHIQGQGLSGQGLCARALYDYQAADDTEISFDPENLITGIEVIDEGWWRG
YGPDGHFGMFPANYVELIE

Human HIP-55 protein sequence - var4 (public gi: 10441970) (SEQ ID NO: 393)
MEKVAKASGANYSFHKESGRFQDVGPQAPVGSVYQKTNAVSEIKRVGKDSFWAKAEKEEENRRLEEKRRA
EEAQRQLEQERRERELREAARREQRYQEQGGEASPQRTWEQQQEVVSRNRNEQESAVHPREIFKQKERAM
STTSISSPQPGKLRSPFLQKQLTQPETHFGREPAAAISRPRADLPAEEPAPSTPPCLVQAEEEAVYEEPP
EQETFYEQPPLVQQQGAGSEHIDHHIQGQGLSGQGLCARALYDYQAADDTEISFDPENLITGIEVIDEGW
WRGYGPDGHFGMFPANYVELIE

Human HIP-55 protein sequence - var5 (public gi: 10121215) (SEQ ID NO: 394)
MAANLSRNGPALQEAYVRVVTEKSPTDWALFTYEGNSNDIRVAGTGEGGLEEMVEELNSGKVMYAFCRVK
DPNSGLPKFVLINWTGEGVNDVRKGACSSHVSTMASFLKGAHVTINARAEEDVEPECIMEKVAKASGANY
SFHKESGRFQDVGPQAPVGSVYQKTNAVSEIKRVGKDSFWAKAEKEEENRRLEEKRRAEEAQRQLEQERR
ERELREAARREQRYQEQGGEASPQSTWEQQQEVVSRNRNEQESAVHPREIFKQKERAMSTTSISSPQPGK
LRSPFLQKQLTQPETHFGREPAAAISRPRADLPAEEPAPSTPPCLVQAEEEAVYEEPPEQETFYEQPPLV
QQQGAGSEHIDHHIQGQGLSGQGLCARALYDYQAADDTEISFDPENLITGIEVIDEGWWRGYGPDGHFGM
FPANYVELIE